THE INFLUENCE OF RECOMBINATION ON THE DIVERSIFICATION OF THE MURINE I REGION

By

ROY WILLIAM TARNUZZER

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1988

This dissertation is dedicated to Lena
as a token of my appreciation for
the love, support, and encouragement
she has provided.

ACKNOWLEDGEMENTS

I would like to thank my mentor and committee chairman, Dr. Edward K. Wakeland, for his input and help in the compilation of this dissertation.

I would also like to express my appreciation to my other committee members, Drs. Ammon Peck, Bill Winter, Harry Ostrer, and Linda Smith, for their help and interest in this research.

My fellow students and laboratory technicians in the lab definitely have made these past five and a half years quite interesting. I would especially like to thank Stefen Boehme for the "unique" environment we were able to maintain in the lab and the many entertaining excursions we made into the outside world. And of course, what would life be without Rick McIndoe, the only human being who could put up with anything I could dish out. I would also like to extend a thank you to all the other people in the lab who, through the years, have helped and encouraged me.

I would also like to thank the many other people in the department for anything they might have done to make my stay in the Department of Pathology a pleasant one.

Finally, I would like to thank all those fish I have caught over the last half a dozen years for the healing therapy they provided when it was needed most. Thanks are extended to my brother, Bob, for without him and the

numerous fishing trips, all those fish would have gone through life without thrill of fighting me at the other end of the line.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

THE INFLUENCE OF RECOMBINATION ON THE DIVERSIFICATION OF THE MURINE I REGION

By

Roy William Tarnuzzer

December 1988

Chairman: Edward K. Wakeland
Major Department: Pathology and Laboratory Medicine

Seven single copy DNA probes were isolated that span 110 kilobases of the Murine I region and used in a Restriction Fragment Length Polymorphism (RFLP) analysis with five restriction endonucleases on genomic DNAs from 28 H-2 homozygous wild and laboratory inbred mice. Polymorphic restriction sites were used to cluster alleles at each locus to form lineages representing groups of minor variants of a single progenitor allele. These lineages were then used to identify recombinational events between the loci probed. Three recombinational hotspots (RHS) were identified from the 26 unique I region haplotypes analyzed. These RHSs are located; 1) in the second intron of Eβ, 2) at the centromeric end of Eα, and 3) approximately 5kb telomeric of the Aα gene. The Eβ and Eα RHSs correspond to those already documented while the RHS adjacent to Aα has not been previously defined. This RHS maps to a 4.7 kb stretch of DNA 3' of the Aα gene and

its activity appears to be haplotype dependent. The three RHSs separate the I region into four genomic segments where the sequences within a particular segment accumulate mutations at the same rate. These segments were termed recombinationally depressed segments (REDS) since recombination is localized to the RHSs with only a few rare recombinational events occurring within a defined REDS. These REDS were grouped into lineages which represent a limited number of evolutionary units which are shuffled between haplotypes during evolution. The genes within a REDS, for example, Aα and Aβ in REDS1, show a strong linkage disequilibrium which results in the coordinate evolution of these two genes. In the case of the A molecule, this linkage disequilibrium between these co-expressed genes appears to be necessary for the proper expression on the cell surface. This same pattern of evolution is seen in the t haplotype mice which contain a large number of wild type alleles suggesting a much higher degree of recombination between these two different forms of chromosome 17 than previously expected.

CHAPTER I
INTRODUCTION

The murine major histocompatibility complex (MHC) located on chromosome 17 and called H-2 is a multigene family coding for polymorphic surface glycoproteins involved in cell recognition and the generation of immune responses to foreign antigen. The I region, spanning approximately 120 kilobases of DNA, lies within H-2 and encodes the two class II molecules, I-A and I-E. These class II or Ia molecules are heterodimers composed of an α and β chain which non-covalently associate in the cytoplasm and are expressed predominately on B lymphocytes and activated macrophage (Flavell and Widera 1986).

The genes for the class II molecules are arranged from the centromere in the order Aβ, Aα, Eβ, Eβ2, and Eα. The Aβ, Aα, Eβ, and to a lesser extent, Eα molecules, are very polymorphic and show many distinct forms or alleles at the protein and DNA level. The set of alleles present in the I region of a specific mouse constitutes its haplotype.

The high amount of polymorphism, i.e. the large number of distinct alleles, makes the H-2 ideal for the study of homologous recombination and its influence on the evolution of the regions adjacent to the sites of recombination. Homologous recombination is a mechanism by which homologous nucleotide sequences or allelic sequences

on homologous chromosomes are exchanged with high fidelity during meiosis. Homologous recombination can occur anywhere along a chromosome. It has been observed that recombination frequencies can vary for different stretches of DNA of the same length and that genetic map distances do not always agree with molecular map distances (Steinmetz et al. 1982b). This suggests that there are regions of DNA that either concentrate or suppress recombinational events. A site where recombination appears to be localized, first identified in procaryotes, has been termed a recombinational hotspot (RHS) (Song 1985).

The H-2 contains four documented RHSs with two falling within the I region (Steimetz et al. 1987). The RHS in Eβ, defined by 12 breakpoints, is localized to a 10kb stretch of DNA and the RHS in Eα is defined by 7 breakpoints localized to a 12-14kb stretch just centromeric to the gene (Steinmetz et al. 1982b; Lafuse et al. 1986). All RHSs in H-2 show three characteristics in common: 1) high frequency of homologous recombination, 2) localization to a small stretch of DNA, and 3) haplotype specificity (Steinmetz et al. 1987).

The presence or absence of an active RHS in different individuals would be expected to have a distinct influence on the generation of haplotypes over an evolutionary timespan. Homologous equal recombination would shuffle and generate new combinations of alleles which would lead

to new haplotypes in a population. Depending on the number of active RHSs in a population, the extent of allele shuffling would vary as would the number of unique haplotypes. Because recombination appears to be localized to specific sites within the I region, markers located in these regions flanked by RHSs should show linkage disequilibrium.

The aims of this dissertation are to survey a large collection of independently derived I region haplotypes and to identify and localize RHSs. Once characterized, I wanted to determine the influence of these RHSs on the generation of I region haplotypes and the relationships of the genes flanked by RHSs.

CHAPTER II
REVIEW OF THE LITERATURE

The major histocompatibility complex (MHC), located on chromosome 17 in the mouse, was first characterized based on its involvement in graft rejection between different inbred mouse lines (Little and Tyzzer 1916). With the advent of serologic techniques and their application to the study of the H-2 complex (Gorer 1936), the genetics of this region began to interest more and more biologists. From this flourishing interest and advances in various chemical and molecular techniques, a more exact picture of the organization, structure and function of the H-2 and its gene products has emerged.

Organization of the Major Histocompatibility Complex Genomic Characteristics and Structure of Encoded Products

The murine major histocompatibility complex, commonly referred to as H-2, is a large multigene family which codes for the cell surface glycoproteins involved in cell recognition and the control of immune responses to foreign antigens. The H-2 complex, in genetic map distances, encompasses approximately 2 centiMorgans of DNA (Klein 1975) which translates into a physical distance of 2000 to 4000 kilobases (Hood et al. 1982). The H-2 encodes three

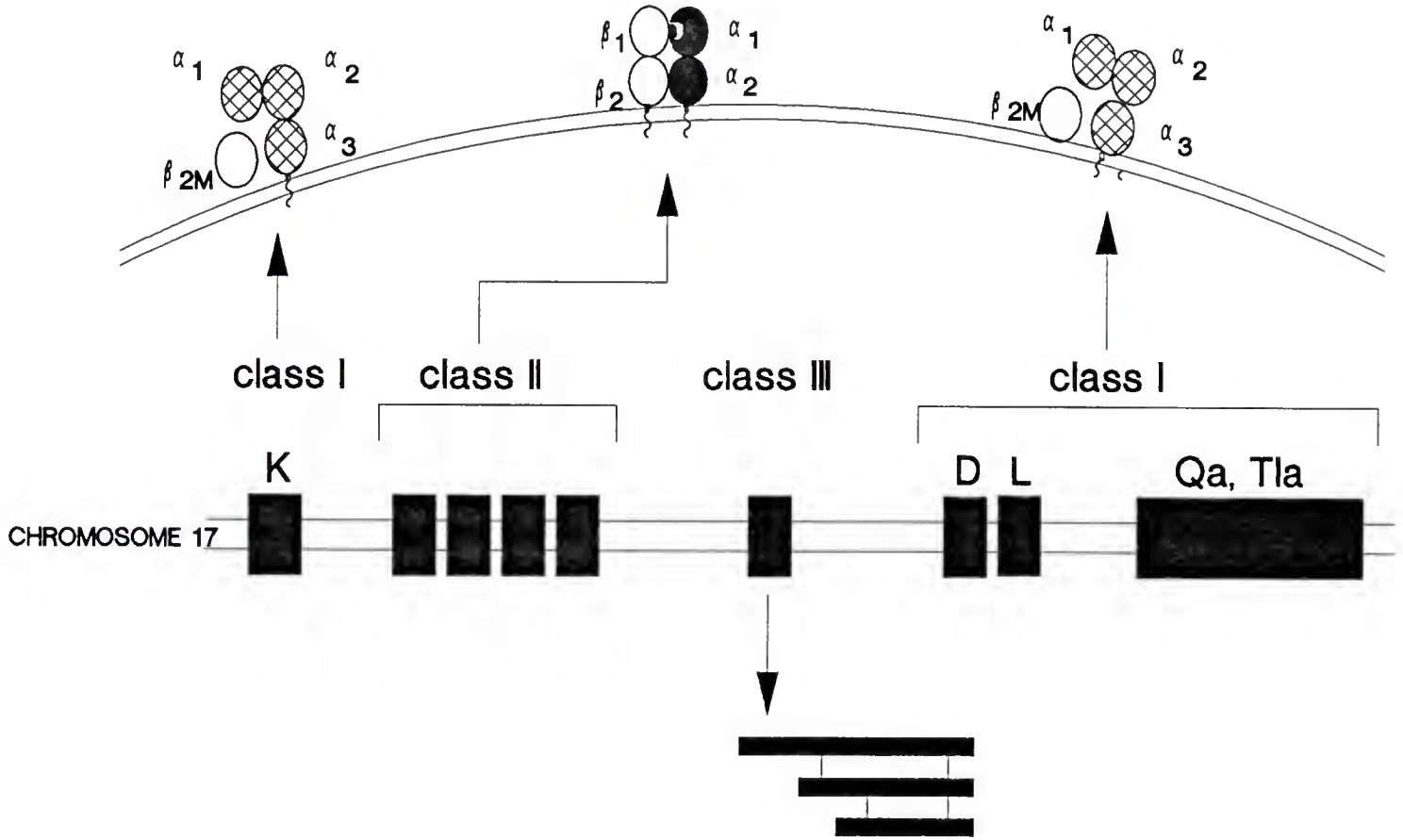
classes of immune related proteins: class I, class II and class III (Klein 1975). Based on functional parameters, the H-2 has been divided into four regions which correspond to the classes of molecules which they encode. The K and D regions contain the class I genes, the I region which contains the class II genes, and the S region contains the class III genes. The class I products fall into 2 general catagories, those involved in graft rejection and those related to development. The first group, the classical transplantation antigens, is encoded by genes denoted K, D, L and R and is expressed on the surface of all nucleated cells. Although they mediate heterologous graft rejection in the laboratory, this is not their function in vivo. These class I molecules function in the restricted presentation of viral and tumor antigens to cytotoxic T lymphocytes (Zinkernagel 1979). The second group of class I molecules is of two families. The Qa family of molecules is expressed on mammalian nucleated blood cells and the Tla molecules are expressed on certain leukemias (Michaelson et al. 1983). Whereas the classical transplantation antigens are very polymorphic, the Qa and Tla antigens exhibit very low polymorphism and their functions are still unknown (Flaherty 1980). Molecular cloning and analysis of the H-2 have revealed over 32 genes spanning at least 800 bps for the Qa and Tla products alone (Steinmetz et al. 1982a). The class I molecules show a unified structure of

three extracellular domains, a transmembrane domain and a cytoplasmic domain which constitutes a 40-45,000 dalton glycoprotein of approximately 350 amino acids. This glycoprotein chain non-covalently associates with a 12,000 dalton molecule encoded on chromosome 2 known as β_2 microglobulin (Klein et al. 1983b).

The class III genes, contained within the S region, encode the complement proteins C2, Bf, Slp, and C4. Although these genes are physically contained in the H-2, Klein and Figueroa (1981) argue against their inclusion in the MHC because they are not functionally related to the class I or class II loci.

The class II genes are contained within the I region, or immune response region, which was first defined by the differential ability of inbred mouse strains to mount an immune response to certain antigens (McDevitt and Sela 1965; Martin et al. 1971) and were latter mapped by the use of recombinant and congenic strains of mice (Benacerraf and McDevitt 1972). There are two class II molecules encoded within the I region, I-A and I-E, assembled from four functional class II genes. These genes are Aβ, Aα, Eβ and Eα as well as the pseudogenes Aβ3, Aβ2 and Eβ2 (Widera and Flavell 1985). A molecular map of the H-2 and the I region is given in Figure 2-1. These class II molecules are composed of a 35,000 dalton α and 29,000 dalton β chain of about 220 and 230 amino acids respectively (Klein et al. 1983b) which non-covalently

Figure 2-1. Diagram of the class I, class II, and class III gene loci and gene products within the major histocompatibility complex located on chromosome 17 of the mouse.


α 1
α 2
α 3
β 2M
β 1
β 2
α 1
α 2
α 1
α 2
β 2M
α 3
class I
class II
class III
class I
K
D L
Qa, Tla
CHROMOSOME 17

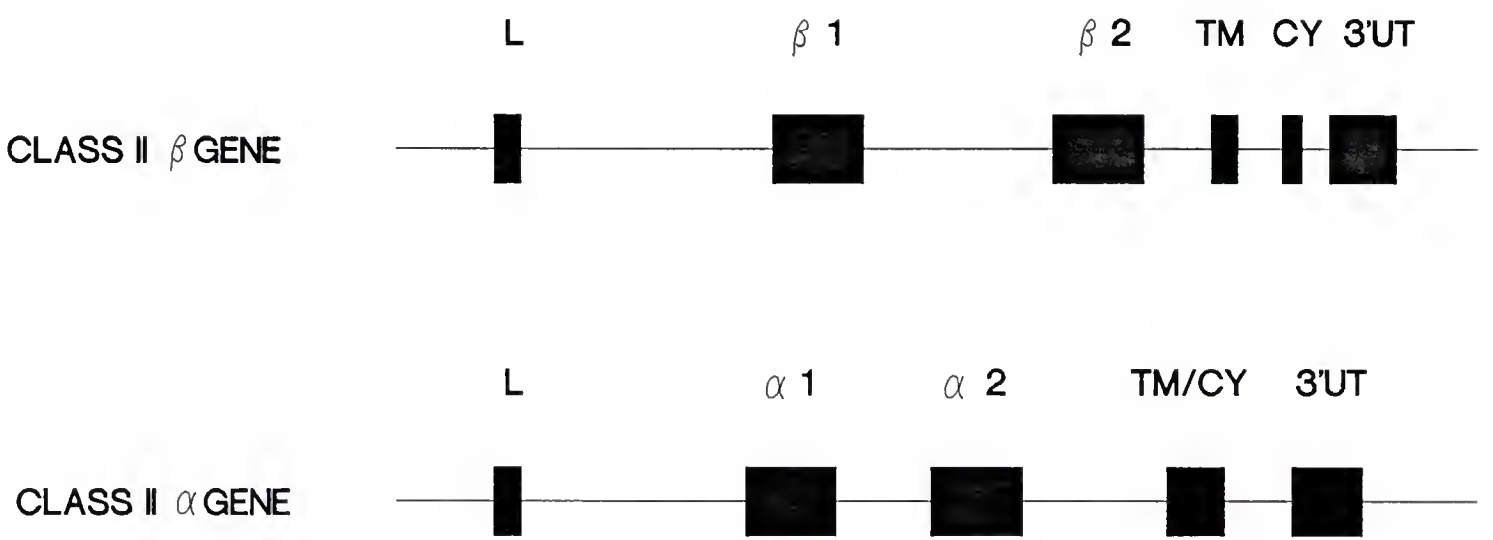
associate in the cytoplasm and are subsequently expressed on the surface of the cell as a heterodimer. The α and β chains are organized similarly into five protein domains. They consist of a hydrophobic leader peptide of 25 amino acids, two 90 amino acid extracellular domains ($\alpha 1 \alpha 2$, $\beta 1 \beta 2$), a hydrophobic transmembrane segment of 25 amino acids and a cytoplasmic domain. The domain structures of the $\alpha 2$, $\beta 1$ and $\beta 2$ regions are formed due to disulfide bonds between pairs of cystine residues located within each domain. The domain organization of the protein directly reflects the intron/exon organization of their respective genes.

The β chain genes of the class II molecules are composed of six exons, one for each protein domain, and an exon for the 3' untranslated region (Saito et al. 1983). The α genes are very similar except they are composed of five, instead of six exons due to the transmembrane and cytoplasmic regions being combined in a single exon (Mathis et al. 1983; McNicholas et al. 1982). A diagram of the organization of the class II α and β genes is given in Figure 2-2.

The Inclusion of the MHC Genes Into the Immunoglobulin Supergene Family

With the advent of cloning and sequencing techniques, a very detailed analysis of class I and class II gene

Figure 2-2. Comparison of the intron/exon organization of the class II alpha and beta genes.


L
β 1
β 2
TM
CY
3'UT
CLASS II β GENE
L
α 1
α 2
TM/CY
3'UT
CLASS II α GENE

structure could be performed. Comparisons of protein and DNA sequences reveal a very similar domain structure for most of the genes of the immune system, with the domain organization reflecting the intron/exon organization of the genes which encode them. This is true for the class II, class I, Thy-1, β2-microglobulin, T4, T8, T cell receptor and immunoglobulin genes (Kaufman et al. 1984; Benoist et al. 1983; McNicholas et al. 1982; Parnes and Seidman 1982; Larhammar et al. 1982; Sukhatme et al. 1985; Maddon et al. 1985; Hood et al. 1983; Davis 1985). The domain and sequence homology within the membrane proximal domains among the genes of the immune system has led to the theory that these genes arose from a single ancestral gene through gene duplication. The strong similarity between the membrane proximal domains of the molecules of the immune system based on size and structure argues strongly for the divergent evolution of a single ancestral gene following gene duplication events (Hood et al. 1983).

I Region Organization

Before the advent of molecular cloning, the I region was considered by immunologists to consist of four subregions as determined by recombinational analysis based on serologic and immune response assays (Klein 1975; Klein et al. 1983a; Mengle-Gaw and McDevitt 1985). The four defined subregions were I-A, I-B, I-J, and I-E. The I-A and I-E subregions were serologically defined and encode

the conventional Ia antigens. The genes for Aβ, Aα, and Eβ map to the I-A subregion, whereas Eα maps to the I-E subregion (Jones et al. 1978; Murphy et al. 1980). The I-B subregion was defined by the regulation of immune responses to IgG_{2a} and lactate dehydrogenase (Lieberman et al. 1972; Melchers et al. 1973). The I-B subregion later became defunct as shown by Dorf and Benacerraf (1975) by the explanation of this immune response phenotype being controlled by the complimentation of two genes, one from the I-A and I-E subregion, respectively. The I-J subregion was defined serologically by reagents directed against an I-J polypeptide, which was believed to be a suppressor factor from suppressor T lymphocytes (Murphy et al. 1976; Murphy et al. 1980). However, attempts to isolate and purify I-J in sufficient quanity for protein sequence analysis have failed. Molecular characterization of the I region by Steinmetz et al. (1982b) in the recombinant strains used to define I-J showed that the product of I-J would have to be encoded by a 3.4 kilobase stretch of DNA. Sequence analysis of this fragment showed that the was no gene which could code for the I-J product (Kobori et al. 1986).

The exact order and number of the class II genes came into view when 240,000 contiguous base pairs of the I region were cloned from the BALB/c mouse (Steinmetz et al. 1982b). Four class II genes were identified with one being a pseudogene due to the lack of hybridization with a

5' probe. It was determined that the BALB/c genome contains two α genes and from four to six β genes. This was confirmed in latter work by Widera and Flavell (1985). The positions of the genes for Aβ, Aα, Eβ, Eβ2, and Eα were conclusively mapped within the I region.

Subsequently, two other class II β genes were discovered and determined to be pseudogenes. Larhammar et al. (1983a) identified Aβ2 and positioned it approximately 20 kilobases centromeric to Aβ. The Aβ2 gene was sequenced (Larhammar et al. 1983b) and the exon/intron organization was found to be the same as for the other class II β genes. The Aβ2 molecule, from the predicted amino acid sequence, shows only 56% homology to the other β chains, in contrast to the typical homology of around 80% seen among these β chains. Based on this, Aβ2 was determined to be the most divergent member of the class II β genes. Widera and Flavell (1985) isolated and characterized Aβ3 and localized it to 75 kilobases telomeric of the K region. Steinmetz et al. (1986) were able to link the Aβ3 gene from BALB/c to the rest of the I region thereby providing a continuous 600 kilobase map of the K and I regions. The pseudogene Aβ3 shows strong homology to the other β genes and 83% homology to the human SBβ gene. Whereas the Aβ2 gene is transcribed (Larhammar et al. 1983a) but not expressed on the cell surface due to splicing errors, Aβ3 shows an 8 nucleotide

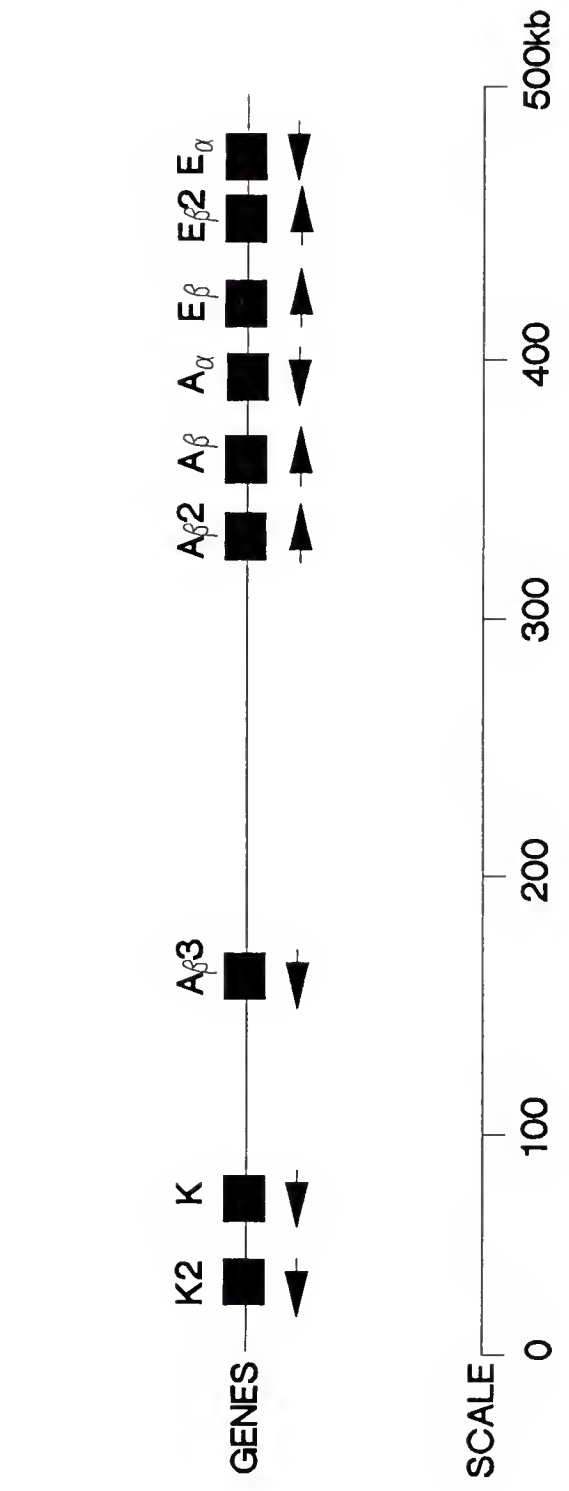
deletion which would make transcription of the gene impossible.

Figure 2-3 illustrates the content and organization of the I region. The genes of the I region are arranged centromerically in the order Aβ3, Aβ2, Aβ, Aα, Eβ, Eβ2 and Eα, and span approximately 300 kilobases of DNA with the functional class II genes confined to a 110 kilobase stretch.

Class II Gene Polymorphism

The MHC genes are the most polymorphic loci known for vertebrates and have made the class I and class II genes of great interest to investigators. Based on serologic and molecular studies, it has been determined that, in general, the Aβ, Aα, and Eβ chains are the most polymorphic (Klein 1975; Benoist et al. 1983), and Eα being the least polymorphic (Klein et al. 1983a). The genes of the class II molecules have been shown to exhibit the same degree of polymorphism as the protein products they encode. This unique variability of the class II genes is therefore a reflection of the unique biological role of these molecules with respect to the immune system.

Figure 2-3. Most recent molecular map of the murine I region. The organization of the class II genes, their transcriptional orientation, and their position relative to the K region are represented. Adapted from Steinmetz et al. (1986).


GENES
K2
K
Aβ3
Aβ2
Aβ
Aα
Eβ
Eβ2
Eα
SCALE
0
100
200
300
400
500kb

Mechanisms for the Generation of Polymorphism

The entire region containing the class II genes has been cloned and analyzed extensively (Larhammar et al. 1983b; Choi et al. 1983; Benoist et al. 1983; Widera and Flavell 1985; Steinmetz et al. 1986). When the I region of several laboratory inbred strains of mice were compared using single copy probes spanning the I region, a variable tract was found in the telomeric half of the I region characterized by extensive sequence diversity determined by an RFLP analysis, and a conserved tract on the centromric end showing very low sequence diversity, with the two tracts meeting at the 3' portion of the Eβ gene (Steinmetz et al. 1984). The genes for Aβ, Aα, and the 5' end of Eβ occupy the 60 kilobases that compose the variable tract. The conserved tract, spanning 50 kilobases, contains the genes for Eβ2 and Eα. Non-coding sequences within the two tracts showed the same patterns of diversity. The mechanisms maintaining this pattern of diversity are not known but perhaps extensive sequence comparisons would shed some light on this unknown.

Nucleotide sequence comparisons of Aβ, Aα, and Eβ reveal extensive sequence polymorphism within laboratory strains of mice (Benoist et al. 1983; Estess et al. 1986). Variations in nucleotide sequence of 5% to 10% are not unusual between alleles of Aβ and Aα with the most diversity localized within the β1 and α1 domains. Choi et al. (1983) did a sequence comparison on genomic clones of

Aβ from the b, d, and k haplotypes and determined that the majority of amino acid substitutions are localized to the animo terminus of the encoded molecule. These mutations are indicative of a pattern of multiple independent events. Recent work by McConnell et al. (1988) on Aβ reveals evidence for segmental exchange between alleles to generate diversity.

Benoist et al. (1983) sequenced six alleles of Aα for the k, d, b, f, u, and q haplotypes and found that most substitutions within the α1 exon are clustered into what they term "regions of allelic hypervariability." A Kabat-Wu variability plot (Kabat et al. 1979) of the corresponding animo acid sequence reveals that amino acid substitutions fall into two hypervariable regions at residues 11-15 and at residues 56-57 which correspond to the regions of the molecule responsible for the binding of foreign antigen (Brown et al. 1988).

Mengle-Gaw and McDevitt (1983) have reported regions of allelic hypervariability between alleles of Eβ also. These regions of diversity are localized to the β1 exon and are separated by tracts of sequence homology which the authors suggest might reflect diversification by gene conversion.

The mechanisms for the generation of diversity of the class II genes are unknown; however, two hypotheses dominate speculations. The first hypothesis proposes that new alleles in a population are generated by

hypermutational mechanisms such as gene conversion or segmental exchange. Segmental exchange or gene conversion was originally defined in fungi (Radding et al. 1978) and is a mechanism by which DNA sequence is copied or transferred to or from genes, usually belonging to multigenic or multiallelic families (Baltimore 1981; Robertson 1982). During meiosis or mitosis, there is pairing of partially homologous sequences followed by mismatch repair thereby converting part of one sequence to that of another. Gene conversion events are characterized by clusters of substitutions at the DNA level. This pattern of diversity is clearly documented in class I genes (Mellor et al. 1983; Weiss et al. 1983) and evidence for the same mechanism of diversification of the class II genes, although to a lesser extent, is also seen (Mengle-Gaw et al. 1984; Widra and Flavell 1985; McConnell et al. 1988).

As mentioned earlier, the majority of mutations within the class II genes appear to be clustered into tracts of allelic hypervariability. Direct evidence for gene conversion in class II genes has been reported by Mengle-Gaw et al. (1984), where an alloreactive T cell clone reacted with determinants present on both $E\beta^b$ and $A\beta^{bm12}$. Sequence comparisons between $A\beta^b$, $A\beta^{bm12}$, and $E\beta^b$ (Choi et al. 1983; McIntyre and Seidman 1984) reveal sequence homology between bm12 and $E\beta^b$ where it differs from $A\beta^b$. The region that is exchanged, encompassing

approximately 14 nucleotides, is flanked by regions of exact homology extending for distances of 20 base pairs either side of the recombinational event.

By examining the nucleotide sequence of eight alleles of Aβ, McConnell et al. (1988) found that, for six of the eight alleles, the evolutionary lineage of the β1 and β2 exons corresponds to the presence or absence of a retroposon insertion within the second intron which is used to define these lineages. The β1 exon of two alleles, Aβ^{b} and Aβ^{nod}, did not reflect their evolutionary lineage by RFLP, and therefore reflects the exchange of sequence, by segmental exchange, from alleles of a different evolutionary lineage.

The second hypothesis for the generation of diversity, termed "trans-species evolution," proposes that the polymorphism arose from the steady accumulation of mutations over long evolutionary periods, and multiple advantageous alleles have survived speciation (Klein 1980). Trans-species evolution, therefore, represents a mechanism for the maintenance of diversity in natural populations. A recent report has shown (McConnell et al. 1988) that 90% of 115 Aβ alleles examined by RFLP analysis fall into two evolutionary lineages based on the presence or absence of a short interspersed nucleotide element (SINE). Using the SINE sequence as an evolutionary tag for the analysis of nine separate species and sub-species of the genus Mus, the authors determined that the SINE

sequence could be identified in species that diverged over eight million years ago. Therefore, these alleles containing the retroposon insertion must have survived speciation suggesting the role of trans-species evolution in the generation of polymorphism seen in modern Mus species.

The above findings indicate that both hypermutational mechanisms and trans-species evolution contributes to the diversity of class II genes. The diversity within class II genes is localized to the regions of the molecule responsible for antigen binding, suggesting that selection for functional diversity in the binding sites of these molecules may maintain these polymorphisms in natural populations. Taken together, this suggests that strong selective pressures play an important role in the maintenance of MHC polymorphism.

Functional Role of MHC Polymorphism

The MHC molecules are involved in cell recognition and generation of the immune response to foreign antigen. The interaction of foreign antigen, the class II molecules, and the T cell receptor determines if an animal can mount an immune response. Therefore, the MHC molecules play a pivotal role in the survival of the animal when challenged by pathogens in their natural environment.

Regulation and Expression of MHC Molecules

Class II molecules are expressed predominantly on two cell types, collectively called antigen presenting cells (APC), and typlified by the macrophage and the B lymphocyte. It is well documented that macrophage and macrophage-like cells play a fundamental role in the induction of immune responses. The interaction of the regulatory T lymphocyte and the APC is under the control of the I region of the MHC, termed MHC restriction, and the ability of the APC to present antigen is dependent on the cell surface expression of a class II molecule (Unanue 1983). It has been demonstrated that the expression of class II is not constitutive in macrophages and can come under both positive and negative control (Steinman et al. 1980; Snyder et al. 1982). In the case of positive control, McNicholas et al. (1982) have shown that factors secreted from mitogen activated spleen cells induce the biosynthesis and cell surface expression of MHC antigens. Subsequent studies have determined this factor to be gamma-interferon (Steeg et al. 1982; King and Jones 1983). When macrophages are incubated with immune interferon, there is a coordinate increase in mRNA for the four class II chains within an eight hour period (Paulnock-King et al. 1985). Further studies on class I induction on macrophages by gamma-interferon suggest the role of a common sequence in the promoter of the genes, in association with a functional enhancer sequence, necessary

for the induction of their expression (Israel et al. 1986).

The B lymphocyte, in contrast to the macrophage, shows a heterologous constitutive level of class II on its cell surface (Mond et al. 1981; Monroe and Cambier 1983). Mitogen activated T cell supernatants were shown to increase the levels of cell surface Ia on resting B cells (Roehm et al. 1984), and later studies identified this factor as B-cell stimulatory factor 1 (BSF-1) (Noelle et al. 1984). BSF-1, or interluekin 4, induces mRNA levels within one hour and cell surface levels as early as two hours (Polla et al. 1986).

These two mechanisms of class II induction reflect the importance of the cell surface expression of class II for the interactions of the regulatory T lymphocytes and the APC for the initiation of an immune response.

Chain Association and the Functional Expression of Class II Molecules

Early studies hava shown that the class II molecules are heterodimeric in nature requiring the association of an α and β chain (Cullen et al. 1974; Jones et al. 1978). By evaluating the functional role of the class II molecules, it was observed that certain immune responses in recombinant mice of the b and k haplotypes mapped to separate subregions of the I region and were therefore under the control of two genes (Jones et al. 1978).

It was noted that some laboratory inbred mice carry mutations that cause the failure of expression of the class II E molecule on the cell surface (Jones et al. 1981). The cloning and analysis of the I region by Steinmetz et al. (1982b) showed that the genes for Eβ and Eα are present in the strains of mice that do not express an E molecule. These defects in expression fall into three catagories (Hyldig-Nielson et al. 1983; Mathis et al. 1983): The H-2^b and H-2^s haplotypes have a deletion in Eα, the H-2^f haplotypes makes an Eα message of abarent size, and the H-2^q haplotype has a defect in the stability of the Eα message. Lack of E molecule expression has been documented to be as high as 30% in wild mice with levels of 50% in the t haplotypes, which can be found in frequencies of up to 40% in wild populations (Nizetic et al. 1984). Eighteen t haplotype strains were shown to lack expression of an E molecule (Dembic et al. 1984). Sixteen of the eighteen strains carry a deletion in the promoter of Eα identical to that seen in inbred mouse strains. The three non-expressing strains which do not carry this deletion carry a mutation where the gene is transcribed but no protein is expressed on the cell surface. These three mutations represent the extreme case where changes in one chain of the class II molecule effect cell surface expression and the ability of an animal to mount an immune response to certain antigens.

There are no reports of the lack of expression of an A molecule within natural populations of mice. The importance of maximizing the amount of class II variation is believed to be reflected in the observation that α and β chains of a given isotype (i.e. A or E) can transassociate in heterozygotes (Fathman and Kimoto 1981). These findings have given rise to the notion of free association among alleles within an isotype. Studies on wild derived haplotypes, by the analysis of tryptic peptide fingerprints from serologically related groups of mice (Wakeland and Klein 1983), show that Aα and Aβ within these strains differ by less than 10% of their tryptic peptides (Wakeland and Darby 1983). RFLP analysis of Aβ and Aα for this same allelic family corroborate this observation at the DNA level (McConnell et al. 1986). Recent studies that indicate that polymorphism can dramatically affect the Aα and Aβ subunits' ability to assemble as an A molecule for functional expression on the cell surface (Germain et al. 1985; Gilfillan et al. 1988). In these studies, Aα and Aβ genes from different haplotypes were either co-transfected or introduced into transgenic mice. It was observed that, haplotype mismatched chains can not effectively associate to get appreciable levels of the transassociated pairs. Taken together, these findings suggest that, in order for proper assembly and cell surface expression, the α and β chains

of the A molecule need to be co-adapted and, therefore, be from the same or similar haplotype.

Role of Class II in the Presentation of Foreign Antigen

It is the interaction of foreign antigen, class II molecules and the T cell receptor which determines if an animal will mount an immune response. Unlike the B lymphocyte, the T cell's receptor cannot bind and recognize free antigen (Moller 1978; Moller 1980). It is, therefore, the role of the class II molecule to present antigen in such a way to enable the T lymphocyte to respond and initiate an immune response.

The T cell receptor must recognize a bimolecular ligand composed of the antigen and the MHC class II molecule (Schwartz 1985). Studies have shown that most T cells recognize non-native forms of the antigen as seen with lysozyme (Adorini et al. 1979), ovalbumin (Shimonokevitz et al. 1983), myoglobin (Streicher et al. 1984), and insulin (Falo et al. 1986). Conversion of antigen from a native to a non-native form is termed antigen processing, and it is performed by APCs which express class II antigens (Allen 1987).

The nature of the interaction of MHC molecules and processed foreign antigen is of great interest for the understanding of the functional role of MHC polymorphism. An early study by Babbitt et al. (1985) revealed that immunogenic peptides from hen egg lysozyme bind

specifically to class II molecules from a responder, but not a non-responder, haplotype. Subsequent studies have focused on the residues responsible for this interaction and the exact nature of the binding between antigen and class II (Buus et al. 1986; Sette et al. 1987; Buus et al. 1987). Recently, the three dimensional structure of a class I molecule was determined by X-ray crystallography (Bjorkman et al. 1987) and because of the significant domain and sequence homologies between class I and class II molecules, Brown et al. (1988) propose a similar model for the class II molecule as determined for class I molecules. The cell-surface portions of each subunit contain two domains (α1, α2, β1, and β2) in which the α1 and β1 domains are postulated to jointly form the binding site which interacts with peptide antigens. The binding site is a groove produced by two parallel alpha helixes which rest atop a platform formed by an eight strand beta pleated sheet. The α1 and β1 domains of the class II molecules are postulated to donate one of the alpha helixes and four strands of the beta pleated sheet each which combine symmetrically to form the binding site. The position of the polymorphic residues within the α1 and β1 domains have been shown to occupy sites within this groove thereby representing contact points between MHC, antigen, and the T cell receptor. This model has not been confirmed as of yet by X-ray crystallography, but conforms

to a variety of structural and functional studies (Allen et al. 1987; Buus et al. 1987; Guillet et al. 1987).

Taken together, these data show the importance of the combinatorial associations of the α and β chains of the class II molecules. The influence of polymorphisms on their association can drastically affect the development of an effective immune response. Maintenance of co-adapted α and β chains will insure the proper assembly and expression of a class II molecule.

Homologous Recombination Within the MHC

The mouse MHC offers a unique opportunity for investigating whether homologous meiotic recombination happens at random or at specific sites within the genome. Many recombinant strains of mice have been characterized, and the advent of molecular cloning has made it possible to map the crossover sites.

Hotspots of Homologous Recombination

The DNA in the genome is not a static structure and undergoes reorganizational events during evolution and development. There is exchange of nucleotide sequences between chromosomes by homologous and non-homologous recombination. Recombination is a mechanism by which DNA sequences are exchanged between homologous and non-homologous chromosomes either during meiosis or mitosis.

There are three types of recombination; non-homologous, homologous equal, and homologous unequal. Non-homologous recombination occurs between sites within distinct structural environments which may be located on the same or different chromosomes. Non-homologous recombination is also referred to as site specific recombination and acts in the differentiation of some prokaryotic and eukaryotic cells. Some specific examples of non-homologous recombination are the integration or excision of bacteriophages and bacterial transposons (Bauer et al. 1984; Calos and Miller 1980), and the rearrangement of immunoglobulin and T cell receptor genes in eukaryotes (Tonegawa 1983; Davis 1985). Non-homologous recombination usually represents mitotic or somatic events.

Homologous recombination, or allelic recombination, occurs between homologous or allelic nucleotide sequences on homologous chromosomes. Homologous recombination can generate new combinations of alleles by the exchange of sequences between homologous chromosomes. Homologous equal recombination breaks and rejoins nucleotide sequences at precisely the same position whereas homologous non-equal recombination cuts and rejoins at different locations on homologous chromosomes leading to the accumulation of duplications and deletions.

Homologous recombination can, theoretically, occur anywhere along the chromosome. It has been recognized that recombination frequencies can vary for different

stretches of DNA of the same size and that genetic map distances do not always agree with molecular map distances (Steinmetz et al. 1982b). Both these observations suggest that recombination is site specific where there are regions that either enhance or suppress recombinational events. A site where recombination is localized has been termed a recombinational hotspot (RHS) (Smith 1983).

Recombinational hotspots were first reported in bacteriophage lambda where mutants arose that grew better in E. coli than the wild type phage due to a recombinational event localized to an eight nucleotide sequence. This sequence, GCTGGTGG, or Chi (for crossover hotspot instigator), enhances recombination leading to better growth in the host bacteria (Smith 1983). The function of the Chi sequence has been determined to be the stimulation of homologous recombination. It exerts its greatest activity within 10 kilobases upstream, and there appears to be a second sequence involved which is located downstream (Smith et al. 1981). Chi sequences are found in E. coli at a frequency of one every five kilobases (Malone et al. 1978). The Chi sequence may therefore represent a molecular basis for recombination.

Homologous recombination has also been documented in eukaryotic systems, such as yeasts, which also serve as a vector system for the study of recombination prone sequences (Song 1985). Recombination has been documented in the human genome within the β-globin gene cluster

(Orkin and Kazazian 1984). By examining haplotype associations of polymorphic restriction endonuclease sites, recombination could be localized to a 9.1 kilobase stretch of DNA located between the δ-globin gene and the first exon of the β-globin gene. Because several haplotypes carry identical mutations in the 5' genomic segment which are found in association with different 3' genomic segments, this site was determined to be a recombinational hotspot.

Recombination Within H-2

Hotspots for homologous recombination have been identified and characterized in the MHC of the mouse. Because of the ability to breed homozygous mouse strains and the large number of distinct alleles for the genes within H-2, researchers have been able to compile an extensive collection of intra-I region recombinant congenic inbred mouse strains. These strains were first identified serologically and functionally by demonstrating the co-expression of two distinct parental epitopes for the A and E molecule in a single offspring (Stimpfling and Durham 1972; Benacerraf and Dorf 1976).

Molecular cloning and characterization of the class II genes made it possible to locate the recombinational breakpoints at the molecular level. Steinmetz and co-workers (1982b) did a molecular characterization of nine intra-I region recombinant strains and found that all the

recombinational events map to a single site within the I region. This site is localized to a 3.4 kilobase region encompassing the second intron of the Eβ gene. These breakpoints were further characterized through southern blot analysis by Kobori et al. (1984), and were localized to only 2.0 kilobases. Sequence analysis of three parental and four I region recombinants reveals that, in three of the recombinants, the recombinational event occurs within a 1 kilobase region of DNA (Kobori et al. 1986). Several subsequent studies have identified more intra-I region recombinants in which the breakpoints map to this RHS (Saha and Cullen 1986a, 1986b; Lafuse and David 1986). In all, there have been 12 breakpoints defined, which are localized within 10 kilobases of DNA spanning the Eβ gene.

Shiroshi and co-workers (1982) examined a congenic mouse strain, B10.MOL-SGR (Mus musculus molossinus) and found a tremendously enhanced frequency of recombination between the K and A locus. Steinmetz et al. (1986) examined a similar mouse, CAS4 (Mus musculus castaneus) which shows a recombination rate as high as 1.5% within this same portion of the genome which encompasses 40 kilobases.

A third recombinational hotspot was identified in another strain of M. m. castaneus, CAS3, which exhibits a recombination rate of 0.6% with breakpoints localized to a 9.5 kilobase stretch of DNA between Aβ3 and Aβ2 (Steinmetz

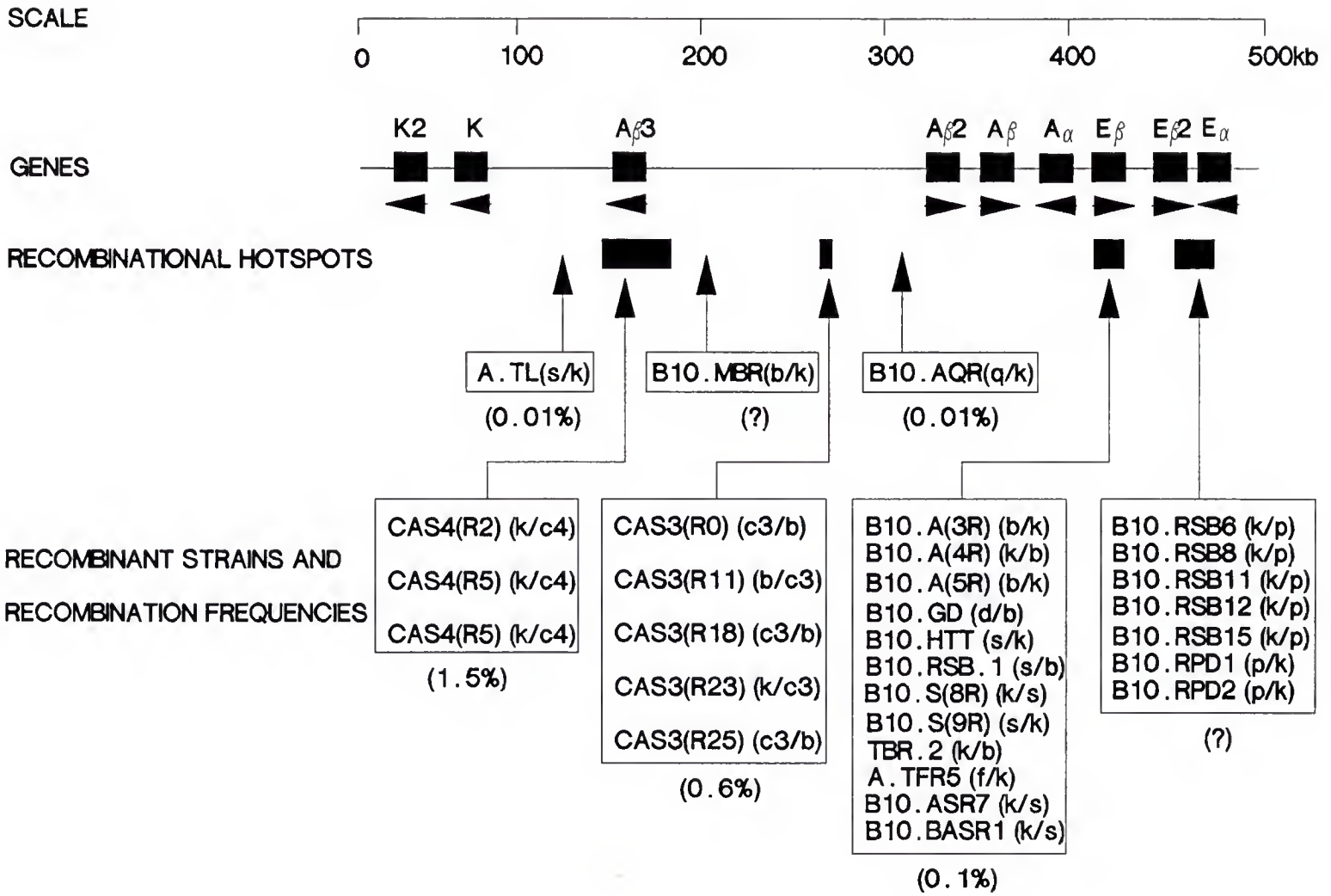
et al. 1986). Further analysis of the nucleotide sequence of five similar recombinant haplotypes revealed that all the breakpoints are confined to a 3.5 kilobase region of DNA (Uematsu et al. 1986). Of the breakpoints examined, all show homologous recombination without any DNA sequences duplicated or deleted between the parental and recombinant haplotypes.

Recently, a fourth RHS has been identified which maps to a 12 to 14 kilobase region centromeric to the Eα gene as characterized by seven breakpoints (Lafuse and David 1986). Figure 2-4 gives a map of the RHSs within the MHC. Earlier data from serologic and tryptic peptide fingerprints (Singh et al. 1981; Wakeland and Darby 1983) have provided evidence for the existence of a fifth RHS within the MHC which maps between the Aα and Eβ genes in some wild derived haplotypes. Data presented in this dissertation help to confirm the existence of this site of recombination.

Haplotype Specificity of Recombination within the MHC

The presence of RHSs at a given site depends on the haplotype of the MHC involved in the genetic cross. Recombination within the hotspot in Eβ is exhibited by strains of the b and k haplotypes in genetic crosses producing recombinant offspring (Steinmetz et al. 1982b; Lafuse et al. 1986). Recombination within two distinct M. m. castaneus haplotypes, c3 and c4, reveal two distinct

Figure 2-4. Location of recombinational hotspots (RHS) within the H-2 complex. Also shown are the recombinant strains used to define each respective RHS, and the recombination frequencies at each site. Adapted from Steinmetz et al. (1987).


SCALE
0
100
200
300
400
500kb
GENES
K2
K
Aβ3
Aβ2
Aβ
Aα
Eβ
Eβ2
Eα
RECOMBINATIONAL HOTSPOTS
A.TL(s/k)
(0.01%)
B10.MBR(b/k)
(?)
B10.AQR(q/k)
(0.01%)
RECOMBINANT STRAINS AND
RECOMBINATION FREQUENCIES
CAS4(R2) (k/c4)
CAS4(R5) (k/c4)
CAS4(R5) (k/c4)
(1.5%)
CAS3(R0) (c3/b)
CAS3(R11) (b/c3)
CAS3(R18) (c3/b)
CAS3(R23) (k/c3)
CAS3(R25) (c3/b)
(0.6%)
B10.A(3R) (b/k)
B10.A(4R) (k/b)
B10.A(5R) (b/k)
B10.GD (d/b)
B10.HTT (s/k)
B10.RSB.1 (s/b)
B10.S(8R) (k/s)
B10.S(9R) (s/k)
TBR.2 (k/b)
A.TFR5 (f/k)
B10.ASR7 (k/s)
B10.BASR1 (k/s)
(0.1%)
B10.RSB6 (k/p)
B10.RSB8 (k/p)
B10.RSB11 (k/p)
B10.RSB12 (k/p)
B10.RSB15 (k/p)
B10.RPD1 (p/k)
B10.RPD2 (p/k)
(?)

recombinational hotspots in the interval between K and Aβ2 (Steinmetz et al. 1986; Uematsu et al. 1986). Recombinational events in the c3 haplotype mapped to the Aβ3/Aβ2 hotspot, whereas recombination in the c4 haplotype occurred in the K/Aβ3 hotspot. Crossing over in the hotspot between Eβ2 and Eα has so far only been detected in crosses of the p haplotype (Lafuse et al. 1986; Lafuse and David 1986).

Molecular Basis For Recombinational Hotspots in the MHC

There are similarities between the Chi sequence in E. coli and the recombinational hotspots in the MHC. As seen for Chi, breakpoints are clustered, only homologous recombination is seen, and the activity of the RHS appears to be dominant. Therefore, the nucleotide sequences around the RHSs were examined for sequences with homology to Chi. Studies by Steinmetz et al. (1986) and Kobori et al. (1986) showed that the region around the Eβ RHS contains a sequence composed of a tetramer, CAGG, which is repeated 22 times. This sequence has limited homology to the Chi sequence of bacteriophage lambda, which is known to promote recombination. A much stronger degree of homology is found between this sequence and the core sequence of human minisatellite DNA, which may facilitate recombination in human chromosomes (Jeffreys et al. 1985). There is no functional evidence, however, to suggest that these repeated sequences are important in recombination

within the MHC. Sequence comparisons can only suggest possible control sequences for recombination, and only functional assays can identify structural elements required for recombination.

Wild Mice

The goals of this dissertation are to survey polymorphism at loci evenly spaced across the I region and to determine the location and influences of recombinational hotspots on the evolution of I region haplotypes in modern species of Mus. Previous studies on this subject were limited in scope due to the nature of the strains of mice used to address these questions. Inbred mice represent only a small subset of highly biased haplotypes. These strains are derived from a limited number of sources that were generated from a high degree of inbreeding, thereby representing a biased sampling of the mouse population and an artificial collection of considerable homogeneity (Ferris et al. 1982; Klein 1974).

Wild mice represent a population whose breeding is not controlled by humans (Bruell 1970), and, therefore, represent a collection of I region haplotypes of considerable heterogeneity. These mice also represent the product of evolutionary processes where the I region haplotypes are fixed and maintained through natural selective pressures.

Natural History of the Wild Mouse

Wild mice can be divided into three groups based on their associations with humans (Sage 1981). Aboriginal mice are free living mice with essentially no interaction with humans. Commensal mice, on the other hand, live in close association with humans, and in most cases rely on humans for their source of food and shelter. The third group, feral mice, represent mice which have made the transition from a commensal association back to the aboriginal state.

The commensal mice fall into four subspecies of Mus musculus; M.m.domesticus, M.m.musculus, M.m.castaneus, and M.m.molossinus (Marshall 1981). The research in this dissertation is concerned only with mice of the M.m.domesticus subspecies. Based on fossil evidence, nuclear genetic variation and mitochondrial genetic variation, Ferris et al. (1983) estimate that this commensal relationship between mouse and man has been in existence for more than 1 million years.

In contrast to aboriginal mice whose range encompasses only the Eurasian continent, commensal mice, also indigenous to Eurasia, have radiated to the new world around much the same time as the first human settlers. The commensal mice have adapted remarkably well to extremely varied climatic conditions with habitats ranging from Europe, the Americas, Australia, Africa, and several South Pacific islands. They may represent the most

evolutionarily advanced member of the genus (Marshall 1981).

M.m.Domesticus is presently found throughout the world, and within its native range, can be found in habitats as diverse as households, agricultural fields, in barren rocky ravines (Gaisler 1975; Hassinger 1973), salt marshes (Breakey 1963), grasslands (Pearson 1963), coal mines (Philip 1938), and mountain enviroments (Harland 1958).

The t Complex

The t complex is a gene complex located on the centromeric one third of chromosome 17 adjacent to H-2, and accounts for nearly 1% of the mouse genome. There are two major structural forms of chromosome 17, the wild type and the t form. The t form is carried in 10% to 40% of wild mice (Artzt et al. 1985; Dembic et al. 1984). A complete t haplotype is one that, by definition, suppresses recombination along the entire 12 centiMorgan region from the gene locus Brachyury (T) to the H-2 complex. The different t haplotypes are all structurally related to one another. Within the chromosomal region occupied by the t haplotypes, there are genes common to non-t bearing mice and mutant genes characteristic of the t complex.

Mutant genes within the t haplotypes have been shown to cause abnormalities in tail length, embryogenesis,

fertility, male transmission ratios, and meiotic recombination (Dunn and Gluecksohn-Schoenheimer 1950; Silver 1985). The t complex has been termed "selfish DNA" (Klein et al. 1986) which serves no apparent purpose other than self propagation and dissemination throughout a population. Two reasons for the prevalence of t chromosomes in the wild are their ability to sway their own transmission and their ability to keep the genetic elements responsible for segregation distortion together by the suppression of recombination.

The molecular nature of the segregation distorters is not known. It is well documented that wild males carrying a complete t haplotype will transmit their t chromosome to greater than 90% of their offspring (Lyon and Meredith 1964a, 1964b). Mice carrying a partial t haplotype can transmit it only when complimented by another chromosome which can restore the transmission distorter (Silver 1985). Lyon (1984) suggests that a series of distorter loci, Tcd, can act on a single responder locus, Tcr. The Tcd loci act in an additive fashion in either a cis or trans configuration to the Tcr, and when the additive effect of the Tcd loci reach a certain level, a high degree of transmission of that chromosome is seen.

The mechanism responsible for the suppression of meiotic recombination is much more straight forward than for transmission distortion. The partial t haplotypes were an important tool in elucidating the molecular basis

of the recombination suppression. The region of suppression in the partial t haplotypes extends only as far as the t DNA present (Bechtol and Lyon 1978; Bennett et al. 1978). Normal levels of recombination are observed between t chromosomes as opposed to the wild type. This suggests that the structure of the t haplotypes are similar to each other, yet different than wild type DNA (Artzt et al. 1982a; Condamine et al. 1983). Subsequently, it was shown that the t haplotypes have a proximal inversion encompassing T and the Tcp (t complex proteins) products (Herrmann et al. 1986), and a distal inversion containing tf (tufted locus) and H-2 (Artzt et al. 1982b; Shin et al. 1983; Shin et al. 1984). Recombination, therefore, is suppressed between t and the wild type due to the inversion of these regions.

The other characteristics of the t complex, sterility and lethality, have been suggested to be secondary add-ons to the primary properties of the t chromosome; transmission distortion and recombination suppression (Klein et al. 1986). This region of the chromosome is believed to carry genes instrumental in embryogenesis and development which have become mutated, hence the lethality and sterility seen, and which are carried along with the "selfish DNA" and disseminated throughout wild populations. This hypothesis is supported by reports that that lethality mutations appear to be single locus mutations which can compliment each other in genetic tests

(Bennett 1975; Klein *et al.* 1984; Winking and Guenet 1978).

Due to the inclusion of the MHC in the recombination suppression of the *t* haplotypes, the association of the alleles at MHC loci with the *t* haplotypes is of great interest. *t* forms of chromosome 17 are believed to be of single founder origin, or at least of a limited founder origin (Klein *et al.* 1986). This suggests that the *H-2* haplotype associated with a *t* haplotype will represent a unique and separate evolutionary lineage than those seen in the wild type. Figueroa *et al.* (1985) have reported the existence of three groups of class II alleles associated with particular *t* haplotypes. Dembic *et al.* (1984) and Nizetic *et al.* (1984) have shown a correlation between a deletion in *E*α and its association with the *t* haplotypes. This identical deletion is seen in wild type forms of chromosome 17 also, which has been interpreted as an ancient origin of this deletion, but which also may be interpreted as having been introduced through recombination with the wild type. A recent study has shown alleles for *A*β shared between *t* and the wild type (McConnell *et al.* 1988). This, in conjunction with data in this dissertation, suggest a higher degree of recombination between the *t* and wild type forms of chromosome 17.

Class II Gene Polymorphisms in Wild Mice

Evidence for the presence of H-2 specificities which are unique to wild mice, and not present in panels of laboratory strains, led to the quest for new alleles from wild mouse populations. Serology was a powerful tool for the analysis and characterization of these H-2 specificities, but there was a problem in separating these reactivities from non-H-2 antigens and other H-2 specificities in the heterozygous animal. This problem led to the development of wild derived congenic mouse lines on a B10 backround, collectively referred to as the B10.W congenic lines (Klein 1973, 1975). Wild male mice were bred with B10.BR female mice and the offspring were backcrossed 8 to 14 times with the continual selection of an H-2 marker specific for the wild haplotype. These lines were maintained by brother x sister matings and the wild H-2 haplotypes selected for on a C57BL/10 background.

Serologic examination of the B10.W lines revealed the extreme polymorphisms of the class II genes within wild populations (Klein 1975; Zaleska-Rutcznska and Klein 1977). Of the 16 haplotypes examined, a few appear identical to inbred haplotypes, a few are identical to each other, but the majority represent novel haplotypes which are not seen in laboratory inbred strains. Later serologic analysis of 29 wild derived haplotypes by Wakeland and Klein (1979a; 1981), revealed three new I region haplotypes; u, v, and j. These same types of

analyses of wild haplotypes show evidence for possible recombination events (Duncan and Klein 1980; Wakeland and Klein 1979b), which points out the value of wild derived H-2 haplotypes for the study of recombination.

Tryptic peptide mapping, in conjuction with the serologic data, demonstrate that many of the haplotypes can be grouped into families of varient alleles (Wakeland and Klein 1983). The nature of the variations between alleles of the A molecule within these families were investigated, and it was found that these alleles differ by less than 10% of their tryptic peptides. Most of these differences are localized in the $\alpha 1$ and $\beta 1$ domains of the A molecule (Wakeland and Darby 1983; Wakeland et al. 1985). Studies at the DNA level confirm the relatedness of these alleles, and have added more insight into the mechanisms which play a role in the generation of diversity of class II molecules (McConnell et al. 1986, 1988).

Wild derived lines have continually shown evidence for recombination within the I region (Duncan and Klein 1980; Wakeland and Klein 1979b; Wakeland and Darby 1983; Singh et al. 1981). A recent study by Soper and co-workers (1988) reveals the prevalence of the RHS in $E\beta$ in the diversification of I region haplotypes within a small panel of wild mice. Singh et al. (1981) demonstrated that recombination commonly occurs between $A\alpha$ and $E\beta$, a fact which has not been borne out at the molecular level until

the work in this dissertation. These data suggest the important role of the wild mouse in the understanding of the evolution of the I region and its application to the study of recombination.

CHAPTER III
MATERIALS AND METHODS

Mice

All mice used in this study were from the mouse colony in the Tumor Biology Unit at the Department of Pathology and Laboratory Medicine, University of Florida, or from our wild mouse colony located at the Animal Care Facility, University of Florida. Strains included in this analysis are listed in Table 3-1. The wild derived mouse strains were maintained by brother x sister matings and are homozygous at the H-2 complex unless otherwise noted. t haplotype mice were supplied by Dr. Joseph Nadeau, Jackson Laboratories, and were maintained as heterozygotes due to the lethality of the t mutations carried by these strains.

Antibody Isolation and Conjugation

Monoclonal antibody, 14.4.4, (Ozato et al. 1980) was produced by injecting 0.5 ml containing 2 x 10^6 hybridoma cells interperitoneally into sub-lethally irradiated, Pristan (Sigma, St. Louis, MO) primed, male BALB/c mice. Ascites fluid was harvested every 2 days for a 1 month period. The ascites fluid was run over protein A-sepharose column (Pharmacia Fine Chemicals, Uppsala) at 50

ml/hr and then washed with PBS until the absorbance at 280 nm was at baseline as determined spectrophotometricly. IgG was eluted from the column with a solution of 0.58% acetic acid, 0.15 M NaCl and the eluate collected in 5 ml fractions. The eluted IgG was dialyzed 18 hours at 4°C against pH 9.3 carbonate/bicarbonate buffer (17.3 g $NaHCO_3$/8.6 g $NaCO_3$ in 1 liter H_2O). Fluorescein isothiocyanate (FITC) (Sigma, St. Louis, MO) was dissolved in dimethyl sulfoxide at a concentration of 1.0 mg/ml and the appropriate amount was added to the IgG and allowed to react at room temperature for 2 hours. This mixture was loaded on a G-75 (Bio-Rad Laboratories, Richmond, CA) column in PBS plus 0.1% NaN_3 and the first colored band was collected and used in subsequent experiments.

Spleen Cell Isolation, Immunostaining, and Flow Cytometric Analysis

Freshly explanted spleens were minced through wire screens to make single cell suspensions. Red blood cells were lysed with an ammonium sulfate solution (0.5% w/v) and washed extensively with PBS. 1 x 10^6 cells were resuspended in PBS plus 0.1% NaN_3 and incubated with a 1:150 dilution of the FITC conjugated antibody for 30 minutes at 4°C. Samples were washed 3 times with PBS and brought up in 0.5 ml for flow cytometry. Cells were passed through a 44 μ nylon mesh filter and then run on a

TABLE 3-1. H-2 Homozygous Wild and Inbred Mice.

Strain of *Mus m. domesticus*	H-2	Geographic Origin
B10	b	old inbred
B10.D2	d	old inbred
B10.M	f	old inbred
B10.WB	j	old inbred
B10.BR	k	old inbred
B10.F	p	old inbred
B10.Q	q	old inbred
B10.RIII	r	old inbred
B10.S	s	old inbred
B10.PL	u	old inbred
B10.SM	v	old inbred
B10.SAA48	w3	Michigan
B10.KEA5	w5	Michigan
B10.CAA2	w11	California
B10.STC77	w14	Michigan
B10.STC90	w15	Michigan
B10.CHA2	w26	Michigan
STU	w34	West Germany
AZROU1	w201	Morocco
FAIYUM3	w206	Egypt
FAIYUM4	w207	Egypt
FAIYUM5	w208	Egypt
JERUSALEM3	w215	Israel
JERUSALEM4	w216	Israel
METKOVIC1	w217	Yugoslavia
METKOVIC2	w218	Yugoslavia
METKOVIC3	w219	Yugoslavia
W12A	w233	Amsterdam
TT6	t6	
C3H.tw5	tw5	
C3H.tw8	tw8	
C3H.tw12	tw12	
C3H.tw32	tw32	
C3H.tw71	tw71	
C3H.tw75	tw75	

FACS II fluorescence activated cell sorter (Becton-Dickinson, Mountain Veiw, CA) at a flow rate of 200-250 cells/second.

Isolation of genomic DNA

Genomic DNA was isolated from liver tissue by a Protease K (Sigma, St. Louis, MO)/SDS method as detailed in Maniatis et al (1982). After 24 hrs of starvation, mice were sacrificed and their livers were surgically removed. Liver tissue was minced with scissors and added to a mortar containing liquid nitrogen and ground to a fine powder. The frozen powder was then added to 40 ml of a TES buffer (10 mM Tris HCl, pH 7.5; 5 mM EDTA, 100 mM NaCl) containing 1% SDS and 0.4 mg/ml protease K. This solution was incubated at 65^{o}C in a water bath for 18 hours. DNA solutions were then extracted three times with Tris equilibrated phenol (pH 7.5), twice with a chloroform/isoamyl alcohol solution (25:1 v/v) and precipitated with 2.5 times the volume of isopropanol. DNA was hooked out of solution with Pasteur pipettes, resuspended in 1.0 ml TE buffer (10 mM Tris HCl, pH 7.5; 1 mM EDTA), and dialyzed against TE buffer. Resulting DNA solutions were quantitated spectophotometricly and electrophoresed on 0.7% agarose gels to confirm their high molecular weight.

Endonuclease Digestion and Agarose Gel Electrophoresis

Aqueous solutions containing 20 ug of high molecular weight DNA were digested with 40 units of one of 5 restriction endonucleases; Bam HI, Bgl II, Eco RI, Pvu II, and Sac I for 18 hours at 37°C under conditions described by the supplier (Bethesda Research Laboratories, Bethesda, MD). Complete digestion was confirmed by removing 10% of the reaction mixture, adding 0.5 ug of phage lambda DNA, incubating an additional 4 hours, and electrophoresing on an agarose gel. Characteristic restriction patterns of phage lambda DNA with a homogeneous smear of genomic DNA were indicative of complete digestion of the DNA mixture. The remaining digested genomic DNA solution was electrophoresed on 0.7% agarose gels for 20 hours at 3 V/cm in a water cooled horizontal electrophoresis apparatus (International Biotechnologies Incorporated, New Haven, CT).

Capillary DNA Transfer and Hybridization

Restriction endonuclease digested DNA was transferred from the gels onto nylon membranes (Zetabind, AMF, Meriden, CT) by the method of Southern (1980). The nylon membranes were dried in a vacuum oven at 80°C for 3 hours and stored at room temperature until hybridization. Membranes were washed in a 0.1% SSC (0.015 M NaCl, 0.0015 M sodium citrate)

solution containing 0.5% SDS at 65°C for 1 hour in a shaking water bath. Prehybridization and hybridization of the filters was performed as described by the supplier (AMF, Meriden, CT). Membranes were hybridized with ^{32}P-labeled DNA probes of specific activity of approximately 2 x 10^8 dpm/ug by nick translation (Bethesda Research Laboratories, Bethesda, MD) for 18 hours at 42°C. Non-specifically bound probe was removed by two successive washes in 0.1% SSC/0.1% SDS at 65°C in a shaking water bath. Membranes were exposed for 2-5 days to XAR-5 X-ray film (Kodak, Rochester, NY) with Lightning Plus intensifying screens (Dupont, Wilmington, DE). After autoradiography, hybridized probes were removed by washing with 0.1% SSC/0.5% SDS at 80°C for 20 minutes and sequentially rehybridized with other labeled probes.

RNA Isolation and Analysis

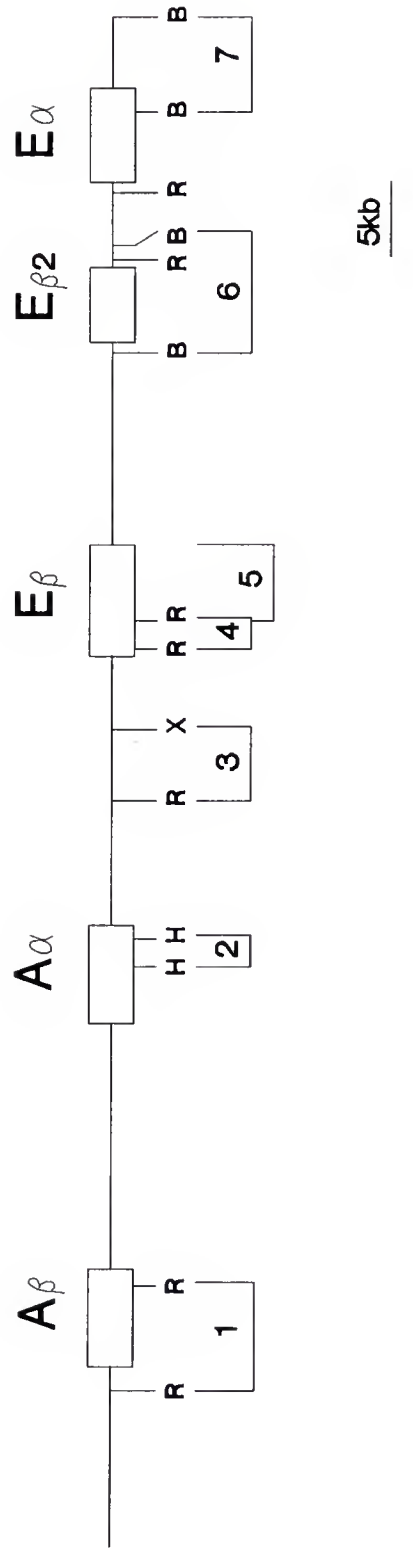
Total cellular RNA was isolated from mouse splenocytes with guanidine isothiocyanate (International Biotechnologies Inc., New Haven CT) as described (Chirgwin et al. 1979; Dingler et al 1986). Isolated spleen cells were dissolved in 4 M guanidine isothiocyanate and layered onto a buffer of 5.7 M cesium chloride (Bethesda Research Laboratories, Bethesda, MD) and centrifuged at 20,000 rpm for 18 hours. The supernatant was aspirated and the precipitated RNA was resuspended in sterile DEPC-treated water, phenol and chloroform extracted twice, and ethanol precipitated. RNA

solutions were stored as precipitates in 100% ethanol at -20°C. The RNA was quantitated spectrophotometrically and 10 ug were electrophoresed in 1.0% formaldehyde\agarose gels to check for degradation. RNAs that were intact were blotted to nylon membranes and hybridized with specific probes as described in the above section. These filters were probed with a 700 bp cDNA for the Ea gene.

Probes

Seven single copy DNA probes were isolated from I region cosmid clones supplied by Dr. Michael Stienmetz, except where otherwise noted, which are evenly spaced across the I region and flanking the known recombinational hotspots. Figure 3-1 shows the location of these probes with relation to the genes within the I region. Probe 1 is a 5.8 kb Eco RI fragment containing the entire $A\beta^d$ gene (Malissen et al. 1983), probe 2 is a 1.2 kb Hind III fragment containing the $\alpha 1$ and $\alpha 2$ exons of $A\alpha^b$ (J. Seidman, personal communication 1984), probe 3 is a 4.2 kb Eco RI/ Xho I fragment midway between $A\alpha$ and $E\beta$, probe 4 is a 1.8 kb Eco RI fragment containing the $\beta 1$ exon of $E\beta^d$, probe 5 is a 700 bp cDNA containing the $\beta 2$, TM, CTY, and 3'UT portions of the $E\beta^d$ gene, probe 6 is a 4.5 kb Bam HI fragment containing the $E\beta 2$ psuedogene, and probe 7 is a 6.5 kb Bam HI fragment containing the 5' promoter region of the $E\alpha^d$ gene.

Figure 3-1. Location of the seven single copy probes used in this research, and their position relative to the class II genes within the I region. Restriction sites which cut out the probes are shown. B= Bam HI, G= Bgl II, E= Eco RI, P= Pvu II, and S= Sac I.


Aβ
R
R
1
Aα
H
H
2
X
R
3
Eβ
R
R
4
5
Eβ2
R
B
B
6
Eα
B
B
R
7
5kb

Genomic Restriction Mapping

Genomic restriction maps were determined for the pertinent strains by first determining the restriction map of the different I region cosmids with the 5 restriction endonucleases used in this study. Likely strains showing recombination were double digested with the restriction enzymes and fragment sizes determined after blotting and hybridization with a probe for the locus of interest. These fragments were then used to map the restriction endonuclease sites for each strain.

Data Analysis

A restriction fragment length polymorphism (RFLP) analysis was performed on the data using equation 21 of Nei and Li (1979) where:

$$F = 2n_{xy} / (n_x + n_y)$$

in which n_x and n_y are the number of fragments from alleles x and y, respectively, and n_{xy} is the number of shared fragments between the alleles. A pairwise correlation analysis of the F values obtained in the RFLP analyses was performed for adjacent loci using a computer program from SAS (Statistical Analysis Software) where a cumulative correlation coefficient (R) was calculated for each allele as compared to all other alleles sampled.

CHAPTER IV
RESULTS

RFLP Analysis of the I Region

An RFLP analysis of the Aβ gene was done previously (McConnell et al. 1986, 1988). Briefly, a 5.5 kb Aβ fragment, probe 1, identified twenty-one alleles with an average F value of 0.29 whithin this panel of mice. This high degree of divergence (low F value) was found to be due to a retroposon insertion within the second intron. This insertion separated the alleles into three evolutionary groups based on either the absence of a retroposon, the presence of a 851 bp retroposon, or the presence of a 1.1 kb retroposon insertion. When these retroposon polymorphisms were taken into account during the RFLP analysis, the mean F value rose to 0.64.

Probe 2, a 1.1 kb Aα fragment containing the α1 and α2 exons, shows a similar high degree of polymorphism as seen with Aβ, and a high number of alleles (Table 4-1). These results show that the degree of diversity detected by each restriction enzyme may vary. By using a combination of several enzymes, the relative diversity between alleles can be better estimated. The data in Table 4-1 also show the frequency of an allele within this panel of mice. For

example, seven strains carry the a allele at this locus whereas only one strain carries the g allele. Another point obtained from these data is the degree of relatedness between particular alleles such that they represent minor variants of each other which differ by only a single restriction fragment, i.e. allele b and allele g by a single Bam HI fragment. The F values calculated between alleles are shown in Table 4-2, and show there are six fragments for each allele for the five enzymes, with each pair of alleles sharing from two to ten fragments. The result of the RFLP analysis is depicted in Figure 4-1, where probe 2 identifies twelve alleles with a mean F value of 0.49 ± 0.18. All polymorphisms at this locus are due to restriction enzyme site changes rather than due to insertions or deletions as determined by restriction fragment length comparisions.

Probe 3, a 4.2 kb fragment midway between Aα and Eβ and designated intergenic sequence 1 (I1), identified thirteen alleles (Table 4-3) in which nineteen of the twenty-eight mouse strains can be grouped into the first four allelic lineages. This probe detects six to seven fragments per allele as seen in Table 4-4, with these alleles showing an average F value of 0.47 ± 0.20 (Figure 4-1). There are no discernable insertion or deletion polymorphisms when the restriction fragment lengths are compared at this locus.

The probe for the 5' portion of the Eβ gene, probe 4, identifies nine alleles (Table 4-5) in which there are only three allelic groups without multiple members, and one

Table 4-1. RFLP sizes and allelic grouping of strains for $A\alpha$.

allele	Bam HI	Bgl II	Eco RI	Pvu II	Sac I	strains
a	b					
a	5.2	3.5	12.0	7.0	13.7 6.6	B10.F, B10.Q, B10.KEA5, MET2, B10.SAA48, B10.CAA2
b	5.4	4.8	12.0	7.0	11.1 6.6	W12A, STU, FAI4, MET3, FAI5
c	5.4	5.2	12.0	7.0	8.0 6.0	B10, JER3, AZR1, B10.STC90
d	5.2	4.8	12.0	7.0	8.0 4.5	B10.WB, JER4
e	5.4	4.8	12.0	4.0	13.7 6.6	MET1, B10.RIII
f	5.4	4.8	12.0	6.0	11.1 4.5	B10.BR, B10.CHA2
g	5.2	4.8	12.0	7.0	11.1 6.6	B10.S
h	5.4	4.8	12.0	7.0	13.7 6.6	B10.SM
i	5.2	5.2	12.0	4.0	10.5 6.0	FAI3
j	5.2	6.5	12.0	7.0	13.7 6.6	B10.PL
k	5.4	3.5	12.0	7.0	8.7 6.6	B10.D2
l	5.2	4.8	12.0	7.0	8.0 6.0	B10.M

[a]Allele designations in ascending alphabetical order based on frequency.
[b]Values are expressed in kilobases.

Table 4-2. RFLP analysis of Aα alleles.

allele	a	b	c	d	e	f	g	h	i	j	k	l
a	-	6/12[b]	4/12	6/12	6/12	2/12	8/12	8/12	4/12	10/12	8/12	6/12
b	.50[a]	-	6/12	6/12	8/12	8/12	10/12	10/12	2/12	6/12	8/12	6/12
c	.33	.50	-	6/12	4/12	4/12	4/12	6/12	6/12	4/12	6/12	8/12
d	.50	.50	.50	-	4/12	6/12	8/12	6/12	4/12	6/12	4/12	10/12
e	.50	.67	.33	.33	-	6/12	6/12	10/12	4/12	6/12	6/12	4/12
f	.17	.67	.33	.50	.50	-	6/12	6/12	2/12	2/12	4/12	4/12
g	.67	.83	.33	.67	.50	.50	-	8/12	4/12	8/12	6/12	8/12
h	.67	.83	.50	.50	.83	.50	.67	-	2/12	8/12	8/12	6/12
i	.33	.17	.50	.33	.33	.17	.33	.17	-	4/12	2/12	6/12
j	.83	.50	.33	.50	.50	.17	.67	.67	.33	-	6/12	6/12
k	.67	.67	.50	.33	.50	.33	.50	.67	.17	.50	-	4/12
l	.50	.50	.67	.83	.33	.33	.67	.50	.50	.50	.33	-

[a]The fraction homologous (F) value as defined by Nei and Li (1979).
[b]Number of shared RF / total RF for both alleles, based on restriction digestion with Bam HI, Bgl II, Eco RI, Pvu II, and Sac I.

Figure 4-1. Molecular map of the I region showing the location of recombinational hotspots (RHS), the number of RFLP alleles distinguished at each locus, and the average diversity (F) calculated among the alleles by an RFLP analysis.

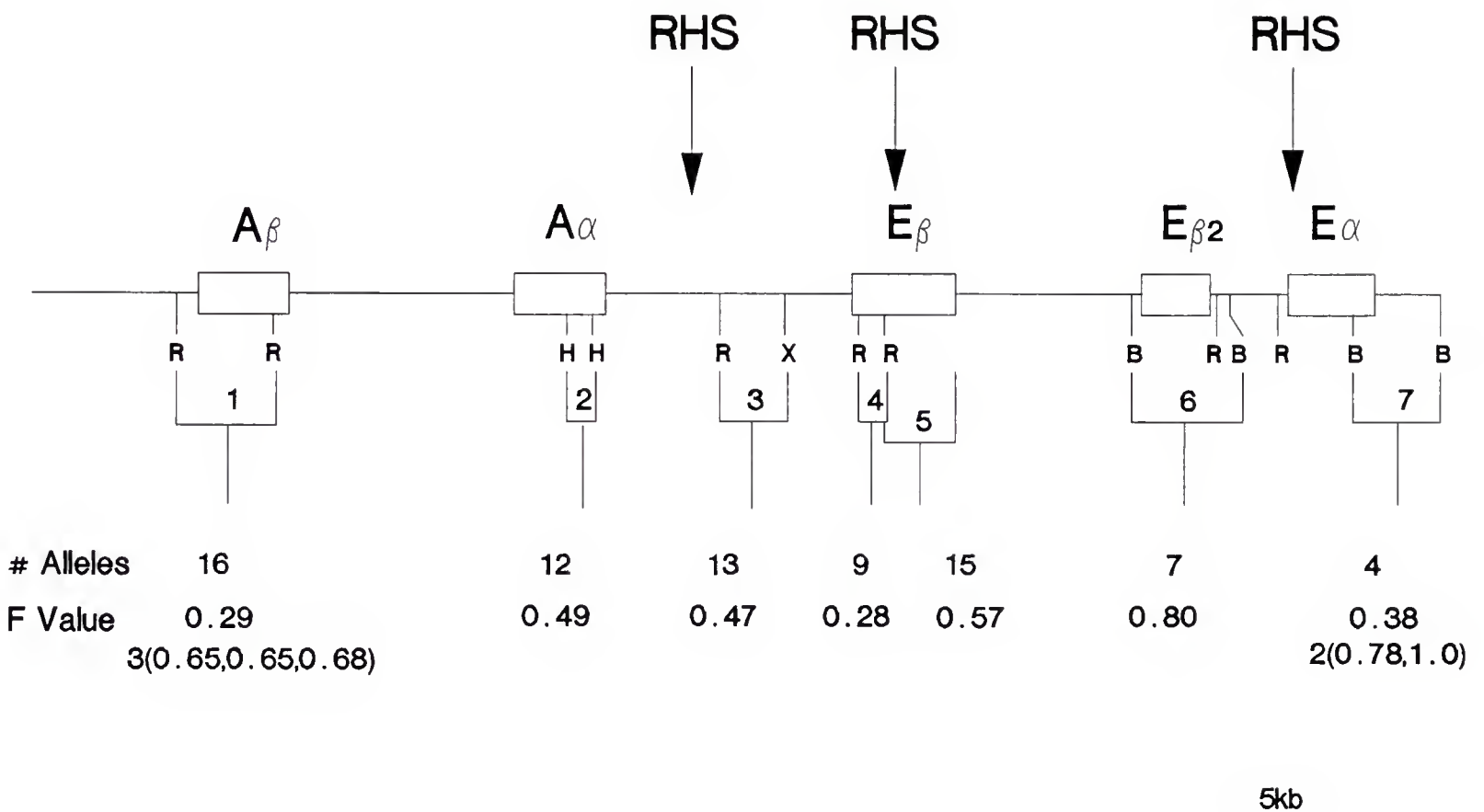

RHS
RHS
RHS
Aβ
Aα
Eβ
Eβ2
Eα
R
R
H H
R
X
R R
B
R B
R
B
B
1
2
3
4
5
6
7
Alleles
16
12
13
9
15
7
4
F Value
0.29
0.49
0.47
0.28
0.57
0.80
0.38
3(0.65,0.65,0.68)
2(0.78,1.0)
5kb

Table 4-3. RFLP sizes and allelic grouping of strains for I1.

allele	Bam HI	Bgl II	Eco RI	Pvu II	Sac I	strains
a	b					
a	10.0	5.0 2.0	9.0 2.0	2.0	9.0	B10, B10.SM, B10.S, AZR1, JER3, MET1, B10.RIII, MET2, B10.STC90, FAI4
b	10.0	5.0 2.0	6.0 2.0	2.0	9.0	B10.Q, B10.CAA2 B10.STC77 B10.CAA2
c	10.0	6.0 2.0	8.5 2.0	6.0	9.0	W12A, STU, FAI5
d	7.0	2.0 0.5	2.0 1.5	2.0	5.0	B10.BR, B10.CHA2
e	9.0	1.5	8.0 2.0	2.0	9.0	MET3
f	6.0	6.0 1.0	6.0 1.0	2.0	8.0	JER4
g	10.0	6.0 2.0	10.0	6.0	9.0	B10.D2
h	7.0	5.0 2.0	8.5 2.0	2.0	9.0	B10.WB
i	10.0	6.0 2.0	10.0 2.0	4.0	9.0	B10.M
j	10.0	5.0 2.0	10.0 2.0	2.0	9.0	B10.SAA48
k	11.0	6.0 2.0	9.0 2.0	2.0	9.0	FAI3
l	10.0	1.5	10.0 2.0	2.0	9.0	B10.F
m	7.0	1.5	9.0 2.0	2.0	9.0	B10.PL

[a]Allele designations in ascending alphabetical order based on frequency.
[b]Values are expressed in kilobases.

Table 4-4. RFLP analysis of I1 alleles.

allele	a	b	c	d	e	f	g	h	i	j	k	l	m
a	-	12/14[b]	8/14	6/14	6/13	2/14	6/13	10/14	8/14	12/14	10/14	8/13	6/13
b	.86[a]	-	8/14	6/14	6/13	4/14	6/13	10/14	8/14	12/14	8/14	8/13	6/13
c	.57	.57	-	4/14	4/13	2/14	10/13	6/14	10/14	8/14	8/14	6/13	4/13
d	.43	.43	.29	-	4/13	2/14	2/13	8/14	4/14	6/14	6/14	4/13	6/13
e	.46	.46	.31	.31	-	2/13	2/12	6/13	4/13	6/13	6/13	8/12	8/12
f	.14	.29	.14	.14	.15	-	2/13	2/14	2/14	2/14	4/14	2/13	2/13
g	.46	.46	.77	.15	.17	.15	-	4/13	10/13	8/13	6/13	6/12	2/12
h	.71	.71	.43	.57	.46	.14	.31	-	6/14	10/14	8/14	6/13	8/13
i	.57	.57	.71	.29	.31	.14	.77	.43	-	10/14	8/14	8/13	4/13
j	.86	.86	.57	.43	.46	.14	.62	.71	.71	-	8/14	10/13	6/13
k	.71	.57	.57	.43	.46	.29	.46	.57	.57	.57	-	6/13	8/13
l	.62	.62	.46	.31	.67	.15	.50	.46	.62	.77	.46	-	8/13
m	.46	.46	.31	.46	.67	.15	.17	.62	.31	.46	.62	.62	-

[a]The fraction homologous (F) value as defined by Nei and Li (1979).
[b]Number of shared RF / total RF for both allele, based on restriction digestion with Bam HI, Bgl II, Eco RI, Pvu II, and Sac I.

Table 4-5. RFLP sizes and allelic grouping of strains for 5'Eβ.

allele	Bam HI	Bgl II	Eco RI	Pvu II	Sac I	strains
a	b					
a	8.6	3.0	2.0	3.4	5.2	B10, B10.S, B10.SM, JER3, B10.STC90, MET1, MET2, FAI4
b	3.8	2.0	4.1	3.4	2.5	B10.F, B10.CAA2, B10.KEA5, B10.Q, B10.STC77
c	4.8	2.7	2.0	3.0	5.0	B10.D2, W12A, STU, FAI5
d	6.9	3.0	2.0	3.4	5.2	B10.BR, AZR1, B10.CHA2
e	3.8	3.0	2.0	3.6	4.2	JER4, B10.WB
f	3.8	3.0	1.0	3.4	2.5	B10.SAA48, FAI3
g	3.8	2.5	2.0	3.4	2.6	MET3
h	3.8	-	17.1	-	-	B10.M
i	8.6	3.0	2.0	3.6	5.2	B10.PL

[a]Allele designations in ascending alphabetical order based on frequency.
[b]Values are expressed in kilobases.

Table 4-6. RFLP analysis of 5'Eβ alleles.

allele	a	b	c	d	e	f	g	h	i
a	-	2/10[b]	2/10	8/10	4/10	4/10	4/10	0/10	8/10
b	.20[a]	-	0/10	2/10	2/10	6/10	4/10	2/10	0/10
c	.20	0	-	2/10	2/10	0/10	2/10	0/10	2/10
d	.80	.20	.20	-	4/10	4/10	4/10	0/10	6/10
e	.40	.20	.20	.40	-	4/10	4/10	2/10	6/10
f	.40	.60	0	.40	.40	-	4/10	2/10	2/10
g	.40	.40	.20	.40	.40	.40	-	2/10	2/10
h	0	.20	0	0	.20	.20	.20	-	0/10
i	.80	0	.20	.60	.60	.20	.20	0	-

[a]The fraction homologous (F) value as defined by Nei and Li (1979).

[b]Number of shared RF / total RF for both alleles, based on restriction digestions with Bam HI, Bgl II, Eco RI, Pvu II, and Sac I.

lineage containing thirteen mice of the H-2^b haplotype. The H-2^p-like alleles show a 1.0 kb deletion when compared to the other alleles in this panel. This deletion was analyzed previously by an RFLP analysis, and was shown to have no effect on the transcription or expression of this gene in the H-2^p haplotype (Soper et al. 1988). Because eight strains share this polymorphism, the average F value for this locus is extremely low, 0.28 ± 0.22 (Figure 4-1 and Table 4-6). There is also one other small 100 to 200 bp deletion detected in the c allele when compared to the a allele for this probe. This deletion also shows no effect on the expression of the gene. When these deletions are taken into account, the average F value rose to 0.60. This locus, and the three previous loci, all map into what was described as the variable tract of the I region (Steinmetz et al. 1984) as characterized by a large number of very distinct alleles.

Probe 5, a 700 bp cDNA for the 3' portion of the Eβ gene, defines fifteen alleles which are more closely related to one another than alleles detected at previous loci (Tables 4-7 and 4-8). Fourteen stains comprise the four most frequent alleles, and ten of the remaining mice represent minor variants differing by a single restriction fragment (RF). Therefore, twenty-four of the twenty-eight strains can be catagorized into these four most common alleles which shared approximately 60% of their restriction

fragments. This homogeneity accounts for the high overall mean F value of 0.57 ± 0.19 (Figure 4-1).

The Eβ2 psuedogene, probe 6, shows seven alleles in which only one of these alleles differs by more than a single RF, and this allele is shared by only three strains (Table 4-9). Table 4-10 demonstrates the relatedness of these alleles as reflected in the extremely high F values between mice. The mean F value for this locus is 0.80 ± 0.08 (Figure 4-1).

The Eα gene is another very non-polymorphic locus which is characterized by probe 7. The Eα gene shows only four alleles which can be grouped into two allelic lineages. Alleles b and c are minor variants of the a allele, one major evolutionary lineage, and allele d-like mice making up the second predominant lineage (Table 4-11). Allele a and its variants represent a closely related family of alleles when compared to the d lineage when the F values generated in an RFLP analysis are examined (Table 4-12). In total, these data show that the Eα alleles can be separated into two classes which correlate to the presence of a 650 bp deletion in the centromeric portion of the gene. The allele a lineage (a, b, and c) do not carry the deletion and, therefore, can transcribe and express this gene. Allele d carrying mice do not have mRNA transcribed and do not express an I-E molecule on their cell surface (Table 4-13). The nature of the defects in E molecule expression within these mice was examined at both the DNA and the RNA level by

Table 4-7. RFLP sizes and allelic grouping of strains for 3'Eβ.

allele	Bam HI	Bgl II	Eco RI	Pvu II	Sac I	strains
a	b					
a	10.0	4.0	12.0	4.5 3.0	2.6 2.4	B10.F, B10.CAA2, B10.STC77, JER4, B10.KEA5
b	10.0	3.1	12.2	4.5 3.0	6.4 2.6	B10.D2, B10.S, FAI5, W12A, FAI5
c	10.0	2.8	12.2	4.5 3.0	6.4 2.5	B10.BR, MET2, FAI4
d	10.0	2.8	6.0	4.5 3.0	6.4 2.5	B10.CHA2, JER3
e	10.0	2.8	12.2	4.5 3.0	2.6 2.4	B10.M, MET1
f	10.0	4.0	12.0	4.5 3.0	6.4 2.6	FAI3, B10.WB
g	10.0	3.1	6.0	4.5 2.7	6.4 2.5	B10.RIII
h	10.0	3.1	6.0	4.5 3.0	6.4 2.5	B10.SM
i	10.0	4.6	12.2	4.5 2.7	6.4 2.6	B10.PL
j	10.0	2.8	12.0	4.5 3.0	2.6 2.4	B10.Q
k	8.6	4.6	12.2	5.0 4.6	2.6 2.4	B10.SAA48
l	10.0	3.1	12.2	4.5 3.0	2.6 2.4	AZR1
m	10.0	2.8	6.0	4.5 2.7	6.4 2.6	B10.STC90
n	10.0	3.1	6.0	4.5 2.7	6.4 2.6	B10
o	10.0	4.6	5.0	4.5 3.0	6.4 2.6	MET3

[a]Allele designations in ascending alphabetical order based on frequency.
[b]Values are expressed in kilobases.

Table 4-8. RFLP analysis of 3'Eβ alleles.

allele	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
a	-	8/14	6/14	6/14	10/14	12/14	4/14	6/14	6/14	12/14	4/14	10/14	6/14	6/14	8/14
b	.57[a]	-	10/14	8/14	10/14	10/14	4/14	6/14	6/14	8/14	4/14	12/14	8/14	10/14	10/14
c	.43	.71	-	12/14	10/14	8/14	8/14	10/14	8/14	8/14	2/14	8/14	8/14	6/14	8/14
d	.43	.57	.86	-	8/14	8/14	10/14	12/14	6/14	8/14	2/14	6/14	10/14	8/14	8/14
e	.71	.71	.71	.57	-	8/14	4/14	6/14	8/14	12/14	6/14	12/14	8/14	6/14	8/14
f	.86	.71	.57	.57	.57	-	6/14	8/14	8/14	10/14	2/14	8/14	8/14	8/14	10/14
g	.29	.29	.57	.71	.29	.43	-	12/14	8/14	4/14	0/14	6/14	10/14	12/14	6/14
h	.43	.43	.71	.86	.43	.57	.86	-	6/14	6/14	0/14	8/14	8/14	10/14	6/14
i	.43	.43	.57	.43	.57	.57	.57	.43	-	6/14	6/14	8/14	10/14	10/14	10/14
j	.86	.57	.57	.57	.86	.71	.29	.43	.43	-	4/14	10/14	8/14	6/14	8/14
k	.29	.29	.14	.14	.43	.14	0	0	.43	.29	-	6/14	2/14	2/14	4/14
l	.71	.86	.57	.43	.86	.57	.43	.57	.57	.71	.43	-	6/14	8/14	8/14
m	.43	.57	.57	.71	.57	.57	.71	.57	.71	.57	.14	.43	-	12/14	8/14
n	.43	.71	.43	.57	.43	.57	.86	.71	.71	.43	.14	.57	.86	-	8/14
o	.57	.71	.57	.57	.57	.71	.43	.43	.71	.57	.29	.57	.57	.57	-

[a]The fraction homologous (F) value as defined by Nei and Li (1979).
[b]Number of shared RF / total RF for both alleles, based on restriction digestion with Bam HI, Bgl II, Eco RI, Pvu II, and Sac I.

Table 4-9. RFLP sizes and allelic grouping of strains for Eβ2.

allele	Bam HI	Bgl II	Eco RI	Pvu II	Sac I	strains
a		b				
a	4.5	4.0 1.0 0.5	3.0 2.5	4.0 3.0	10.0 1.0	B10.D2, W12A, B10.CHA2, B10.S, B10.Q, B10.CAA2, B10.STC77, B10.M B10.PL, STU, JER4, AZR1, FAI3 FAI5, B10.SM, B10.F, B10.KEA5
b	4.5	4.0 1.0 0.5	3.0 2.5	3.5 3.0	10.0 1.0	B10.RIII, MET3, B10.STC90
c	4.5 1.0	4.5 2.5 0.5	3.0 3.8	4.0 1.0	10.0	MET2, FAI4, B10.WB
d	4.5	4.0 1.0 0.5	3.0 2.5	4.0 3.8	10.0 1.0	JER3, MET1
e	4.5	4.2 1.0 0.5	3.0 2.5	3.5 3.0	10.0 1.0	B10
f	4.5	4.0 1.0 0.5	3.0 2.5	4.0 3.0	9.0 1.0	B10.SAA48
g	4.5	4.0 1.0 0.5	3.0 2.5	5.5 4.0	10.0 1.0	B10.BR

[a]Allele designations in ascending alphabetical order based on frequency.
[b]Values are expressed in kilobases.

Table 4-10. RFLP analysis of $E\beta 2$ alleles.

allele	a	b	c	d	e	f	g
a	-	18/20[b]	18/20	16/20	16/20	18/20	18/20
b	.90[a]	-	14/20	16/20	16/20	16/20	16/20
c	.90	.70	-	18/20	14/20	14/20	16/20
d	.80	.80	.90	-	14/20	16/20	18/20
e	.80	.80	.70	.70	-	14/20	14/20
f	.90	.80	.70	.80	.70	-	16/20
g	.90	.80	.80	.90	.70	.80	-

[a]The fraction homologous (F) value as defined by Nei and Li (1979).
[b]Number of shared RF / total RF for both alleles, based on restriction digestion with Bam HI, Bgl II, Eco RI, Pvu II, and Sac I.

Table 4-11. RFLP sizes and allelic grouping of strains for Eα.

allele	Bam HI	Bgl II	Eco RI	Pvu II	Sac I	strains
a	b					
a	6.5	6.0	8.5	6.0	6.0 2.5 1.2	B10.D2, B10.SM, B10.BR, B10.PL, B10.STC90, MET3 B10.CAA2, JER4, B10.SAA48, FAI5 B10.RIII, B10.Q B10.M, B10.KEA5
b	6.5	6.0	8.5	6.0	6.0 1.2 0.7 0.4	W12A, STU, AZR1, JER3, B10.F
c	6.5	6.0	7.5	6.0	6.0 2.5 1.2	B10.STC77, FAI3, B10.WB, B10.CHA2
d	6.0	5.5	8.0	5.5	8.0 2.5	B10, B10.S, MET2, FAI4

[a]Allele designations in ascending alphabetical order based on frequency.
[b]Values are expressed in kilobases.

Table 4-12. RFLP analysis of Eα alleles.

allele	a	b	c	d	
			b		
a	-	12	12	2	
			15	14	13
		a			
b	.80	-	10	0	
				15	15
c	.86	.67	-	2	
					13
d	.15	0	.15	-	

[a]The fraction homologous (F) value as defined by Nei and Li (1979).
[b]Number of shared RF / total RF for both alleles, based on restriction digestion with Bam HI, Bgl II, Eco RI, Pvu II, and Sac I.

Table 4-13. Characterization of E^0 strains.

strain	haplotype	surface expression	Eα message expression	deletion present
AZR1	w201	-	-	-
B10	b	-	-	+
B10.S	s	-	-	+
B10.M	f	-	+[a]	-
B10.Q	q	-	+[b]	-
FAI4	w207	-	-	+
MET2	w218	-	-	+

[a]Mice show abarent sized message.
[b]Mice show low message levels due to a defect affecting the stability of the message.

RFLP and northern blot analysis. Of the three identified defects in Eα gene expression (the 650 bp deletion, a splicing defect, or a message stability defect), only the deletion mutation is found in our panel of wild mice (Table 4-13). This deletion is the only the major polymorphisms seen at this locus and accounts for the low mean F value (0.44 ± 0.38) seen between these mice (Table 4-12 and Figure 4-1).

Probes 5, 6, and 7 all map within the conserved tract of the I-region which is characterized by the low number of very related alleles (Steinmetz et al. 1984).

Allele Lineages

When two or more mouse strains shared an F value of >0.80, thus differing by a single RF, they were grouped into evolutionary lineages. This is demonstrated in Figure 4-2 where, for Aα, there are five strains which carry allele b with two minor variants, bv1 and bv2. Each variant allele is represented by a single member which differs from b by a unique Bam HI and Sac I fragment, respectively. The average F value between mice within a lineage is 0.83 whereas the value is 0.50 between lineages. The number of lineages for each of the loci probed are shown in Table 4-14. Grouping the alleles at each locus across the I region lowers the number of distinct alleles two fold. This helps simplify the data for further analysis and shows that there are a

limited number of old alleles carried by the mice in this panel.

Evidence for Site Specific Recombination Within the I Region: Identification of RHSs in Eβ and Eα

The lineage designations determined in the preceding section were used to determine if and where recombination occurs within the I region haplotypes in our panel. It was expected that if recombination is occurring between two loci, the pattern of allele associations of the recombinant haplotypes will change with respect to that of the donor haplotypes. This is confirmed by the results shown in Table 4-15 when compared to the data presented in Table 4-16. The data presented in Table 4-16 demonstrate the linkage disequilibrium between two loci, I1 and 5'Eβ, seen in the absence of recombination. The results in Table 4-15 demonstrate the predicted switching of allele associations between two loci, 5'Eβ and 3'Eβ, known to undergo recombination. These associations are so pronounced in haplotypes not undergoing recombination that by knowing the lineage of one locus, the lineage of the adjacent locus can be predicted accurately in 90% of the cases. For example, referring to Table 4-16, if a strain belongs to lineage a for I1, then in nine of ten cases these same mice are lineage a for 5'Eβ. This same pattern holds true for all other lineages. In contrast, especially for lineage a, this

Figure 4-2. Grouping of alleles of Aα into evolutionary lineages, a and b, based on the number of shared restriction fragments. Lineages differed by a single restriction fragment. The degree of diversity of alleles within, as contrasted to between, separate evolutionary lineages are also given.

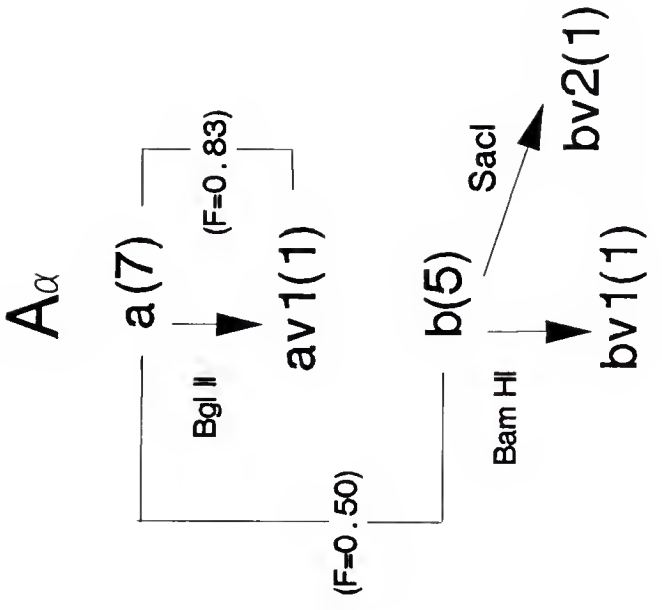

Aα
a(7)
Bgl II
av1(1)
(F=0.83)
(F=0.50)
b(5)
Bam HI
bv1(1)
SacI
bv2(1)

Table 4-14. Allele Lineages Across the I region.

strain	Aβ	Aα	I1	5'Eβ	3'Eβ	Eβ2	Eα
B10	h	c	a	a	gv2	av3	d
B10.BR	p	f	d	d	d	av5	a
B10.CAA2	a	a	av1	b	a	a	a
B10.CHA2	p	f	d	d	dv1	a	av2
B10.D2	cv1	k	cv1	c	b	a	a
B10.F	a	a	av1	b	a	a	av1
B10.KEA5	a	a	av1	b	a	a	a
B10.M	l	dv1	i	h	cv1	a	a
B10.PL	q	av1	m	av1	i	a	a
B10.Q	a	a	av1	b	av1	a	a
B10.RIII	c	e	a	a	av2	av1	a
B10.S	m	bv1	a	a	b	a	d
B10.SAA48	bv1	a	av2	f	k	av4	a
B10.SM	d	bv2	a	a	av3	a	a
B10.STC77	a	a	av1	b	a	a	av2
B10.STC90	n	c	a	a	cv3	av1	a
B10.WB	i	d	h	e	av4	c	av2
AZR1	h	c	a	d	l	a	av1
FAI3	l	i	k	f	av4	a	av2
FAI4	jv1	b	a	a	d	c	d
FAI5	jv1	b	c	c	b	a	a
JER3	h	c	a	a	dv1	av2	av1
JER4	i	d	f	e	a	a	a
MET1	c	e	a	a	cv1	av2	av2
MET2	a	a	a	a	d	c	d
MET3	mv1	b	e	g	o	av1	a
STU	j	b	c	c	b	a	av1
W12A	j	b	c	c	b	a	av1

Table 4-15. Lineage associations within $E\beta$.

strain	5'$E\beta$	3'$E\beta$	recombinant
B10	a	gv2	-
B10.RIII	a	av2	+
B10.S	a	b	+
B10.SM	a	av3	+
B10.STC90	a	cv2	-
FAI4	a	d	+
JER3	a	dv1	+
MET1	a	cv1	-
MET2	a	d	+
B10.PL	av1	i	-
B10.F	b	a	-
B10.KEA5	b	a	-
B10.Q	b	av1	-
B10.STC77	b	a	-
B10.CAA2	b	a	-
B10.D2	c	b	-
FAI5	c	b	-
STU	c	b	-
W12A	c	b	-
B10.BR	d	d	-
B10.CHA2	d	dv1	-

[a]Recombinants defined by switching of lineage association between loci.

Table 4-16. Lineage associations centromeric of $E\beta$.

strain	I1	5'$E\beta$	recombinant
B10	a	a	-
B10.RIII	a	a	-
B10.S	a	a	-
B10.SM	a	a	-
B10.STC90	a	a	-
AZR1	a	d	+
FAI4	a	a	-
JER3	a	a	-
MET1	a	a	-
MET2	a	a	-
B10.CAA2	av1	b	-
B10.F	av1	b	-
B10.KEA5	av1	b	-
B10.Q	av1	b	-
B10.STC77	av1	b	-
B10.SAA48	av2	f	-
FAI5	c	c	-
W12A	c	c	-
STU	c	c	-
B10.D2	cv1	c	-
B10.BR	d	d	-
B10.CHA2	d	d	-

predictive ability is lost due to the high degree of recombination that occurs between 5' and 3'Eβ (Table 4-15). All recombination events at this site occur between an a allele at 5'Eβ and some other allele. Six recombination events are scored at this site as compared to only one possible event between I1 and 5'Eβ in all the mouse haplotypes tested. This is therefore designated a recombinational hotspot (RHS) based on the high frequency of localized recombinational events specific for the a lineage. These recombinational events can be represented graphically (Figure 4-3A), where the fill pattern at each locus represents the lineage origin of the genomic segment of interest. This diagram represents the identification of a RHS within Eβ for mice with the w22, w26, w207, and s haplotypes.

An similar analysis was performed on the genomic segment containing the Eβ2 and Eα genes. Because of the lower degree of polymorphism at the telomeric end of the I region, the 650 bp deletion in Eα is used as a marker in order to identify recombination within this region. Recombination is scored when a deleted allele of Eα is observed adjacent to a proximal expressor associated allele for Eβ2 (Figure 4-3B). Of the three possible lineages of Eβ2 and two lineages at Eα, three recombinant mouse haplotypes can be identified; b, s, and j.

Both of these RHSs correspond to the hotspots previously identified in laboratory inbred strains (Steinmetz et al. 1982b; Lafuse et al. 1986).

Identification of a Recombinational Hotspot Between Aα and Eβ

By the same methods and criteria as above, a third RHS was identified between Aα and Eβ. This recombinational hotspot maps to a 4.7 to 9.2 kb stretch of DNA midway between the two genes (Figure 4-4). The minimum distance was defined at centromeric end by a Bgl II site in w207 and telomeric end by a Pvu II site. These sites were confirmed with a 2.8 kb Bam HI fragment probe which lies equidistantly between the Aα and I1 probes. Five recombinational events were scored at this site and the data for some of the representative haplotypes (s, w207, and w218) are presented in Figure 4-3C. All recombinational events occur between mice of the a lineage for I1, and usually strains of the b lineage for Aα and therefore are haplotype specific. This RHS represents a site for homologous recombination not previously identified in laboratory strains of Mus m. domesticus.

Figure 4-3. Recombination within the I region as determined by the switching of evolutionary lineage of neighboring loci. Each pattern represents the lineage origin of each locus within a haplotype. A) Identification of a RHS in the Eβ gene. B) Identification of RHS in the Eα gene. C) Identification of recombination between the Aα and Eβ genes. Probe location is shown at the top of each diagram, and the haplotype designations are given at the left of each figure.

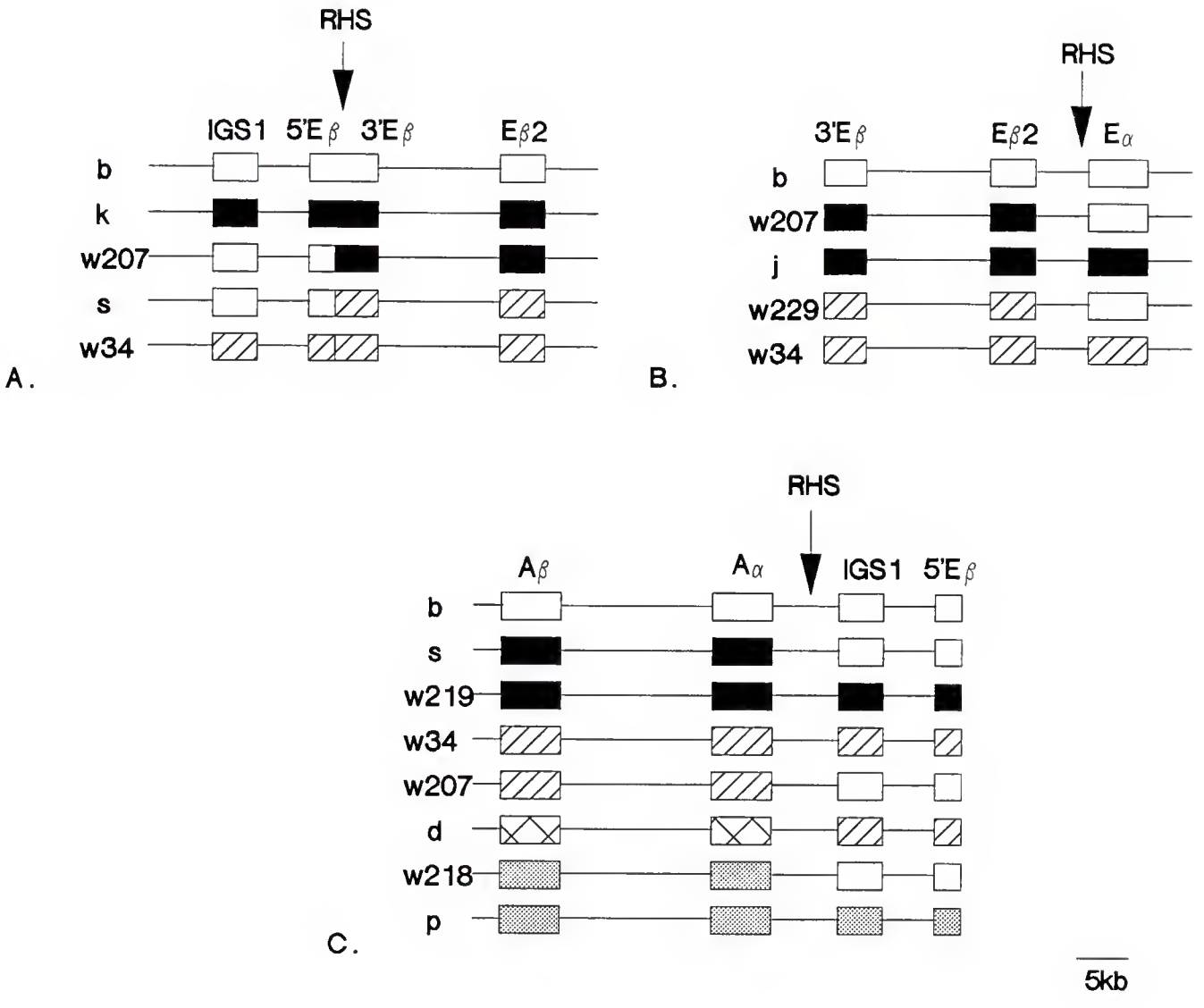
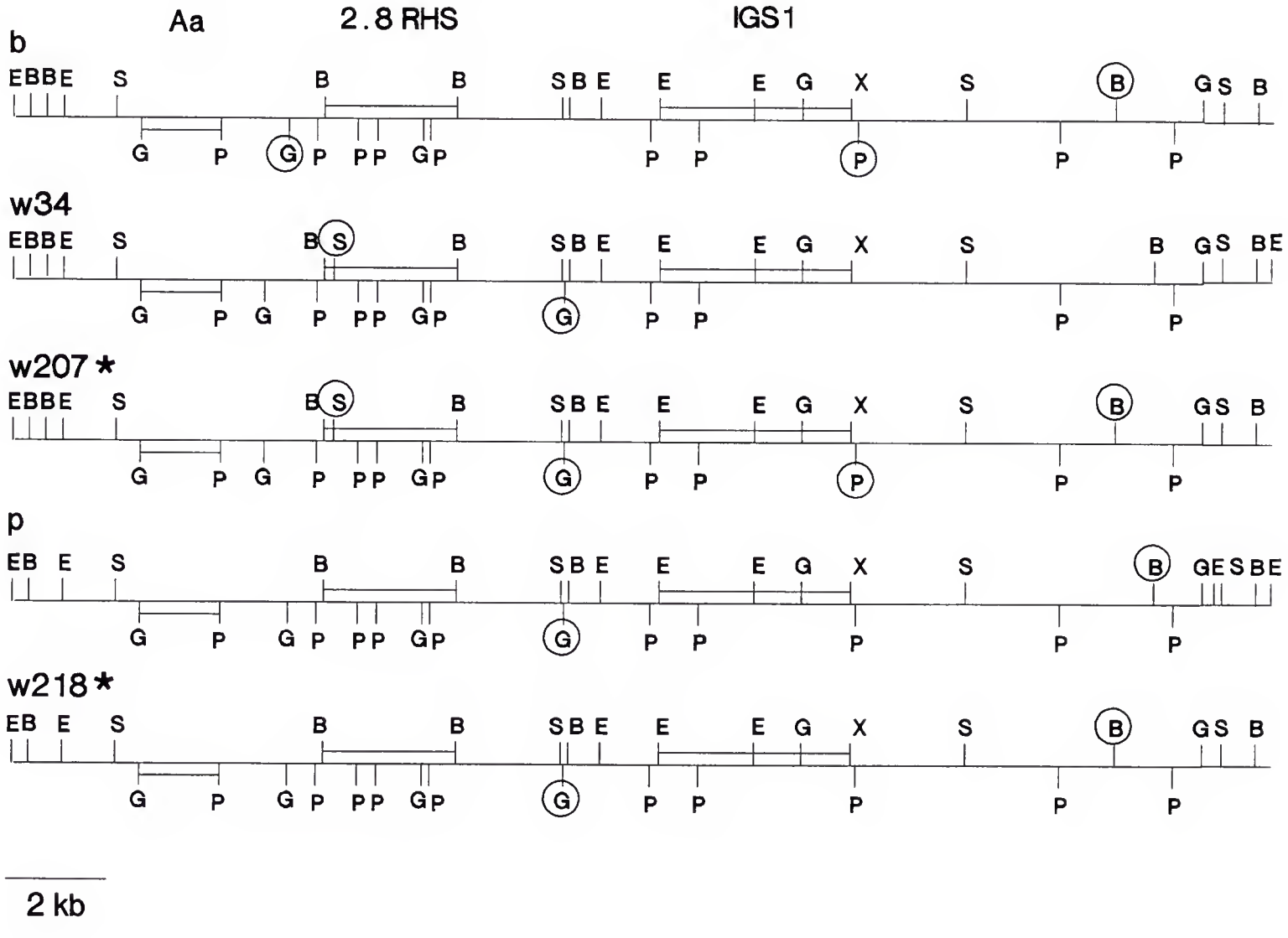

RHS
IGS1
5'Eβ
3'Eβ
Eβ2
b
k
w207
s
w34
A.
RHS
3'Eβ
Eβ2
Eα
b
w207
j
w229
w34
B.
RHS
Aβ
Aα
IGS1
5'Eβ
b
s
w219
w34
w207
d
w218
p
C.
5kb

Figure 4-4. Genomic restriction map of the segment spanning the Aα and Eβ genes within the I region. The location of the probes used to generated these maps are given at the top as well as the sites recognized by the five restriction endonucleases. w207 and w218 represent the recombinant haplotypes identified by this RFLP analysis. B= Bam HI, G= Bgl II, E= Eco RI, P= Pvu II, S= Sac I. The polymorphic restriction sites used to define the bounderies of the recombinational events are circled. Recombinant strains are marked by *.


Aa
2.8 RHS
IGS1
b
EBBE S B B SBE E E G X S B GS B
G P G P PP GP P P P P P
w34
EBBE S B S B SBE E E G X S B GS BE
G P G P PP GP G P P P P
w207*
EBBE S B S B SBE E E G X S B GS B
G P G P PP GP G P P P P P
p
EB E S B B SBE E E G X S B GES BE
G P G P PP GP G P P P P P
w218*
EB E S B B SBE E E G X S B GS B
G P G P PP GP G P P P P
2 kb

Identification of Recombinationally Depressed Segments (REDS) Within the I Region

By lining up the haplotypes in this panel as done for the identification of RHSs, a pattern of associations between loci emerged. It was observed that recombination is localized almost exclusively to the three recombinational hotspots with one event or less occurring elsewhere within the I region. This restricted pattern of recombination at specific, non-random sites separates the I region into discrete genomic segments flanked by the RHSs. There is strong linkage disequilibrium seen between the loci within any of these three genomic segments. Depending on the RHSs active within a particular haplotype, the extent of this linkage disequilibrium varies (Figure 4-5). Some haplotypes carry linkage groups extending across the entire I region, for example H-2^{p}, while others, with more active forms of the different RHSs, show shorter linkage groups. For example, in the H-2^{w208} haplotype, Aβ and Aα form one linkage group while I1 and 5'Eβ form another short linkage group. This pattern of recombination within the I region defined by RHSs, led these genomic segments flanked by RHSs to be defined as recombinationally depressed segments (REDS). These REDS are characterized by the lack of recombination between the loci within a particular REDS and, therefore, a strong linkage disequilibrium is exhibited between these loci. Recombinationally depressed segment 1

contains 44 kb of DNA including the genes for Aβ and Aα, REDS2 contains 19 kb extending just telomeric of Aα to the RHS in Eβ, REDS3 spans 21 kb containing 3'Eβ, Eβ2, and 9kb telomeric to the RHS at Eα, and REDS4 contains Eα and approximately 15 kb telomeric of this gene. These REDS are the genetic units exchanged between haplotypes by homologous recombination during the evolution of the genus Mus which led to the diversification of the I region.

Correlation Analysis of REDS

To test how strong the linkage disequilibrium is between loci within a REDS and across the I region, a statistical pairwise correlation analysis of the F values for each pair of neighboring loci was performed. The correlation coefficient or R values were calculated. These show the relationship between the degree of divergence (F value) for the neighboring loci. For example, by assessing the degree of divergence of one locus, A, in a particular haplotype as compared to a certain subset of haplotypes, if there is correlation between this locus and its neighbor, B, then an analogous pattern of divergence or F value will be seen for locus B as compared to the same subset of haplotypes (Figure 4-6). As the strength of the correlation between loci increases, the R value approaches 1.0. Table 4-17 shows the cummulative results of this type of analysis for all the loci within the I region. The center diagonal

Figure 4-5. Diagram depicting the segmentation of the I region into recombinationally depressed segments (REDS) as influenced by the presence of a RHS. Each pattern represents the lineage for each locus across the I region. Switching of the neighboring patterns indicate recombination leading to the shuffling of the different REDS among I region haplotypes.

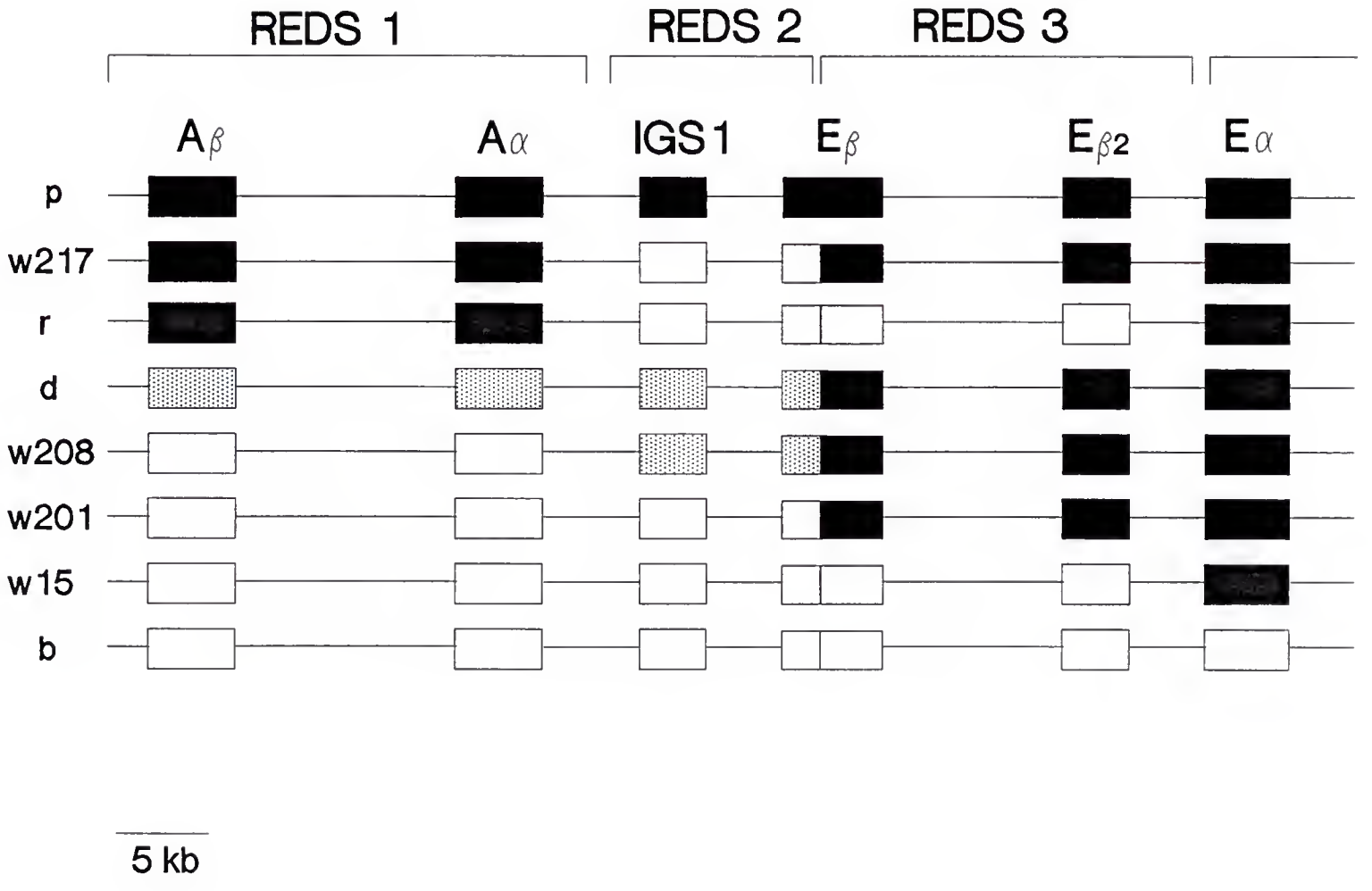

REDS 1
REDS 2
REDS 3
Aβ
Aα
IGS1
Eβ
Eβ2
Eα
p
w217
r
d
w208
w201
w15
b
5 kb

Figure 4-6. Diagram showing the premise behind the correlation analysis of neighboring loci. A positive relationship between the F values from an RFLP analysis of adjacent loci is indicative of strong correlation.

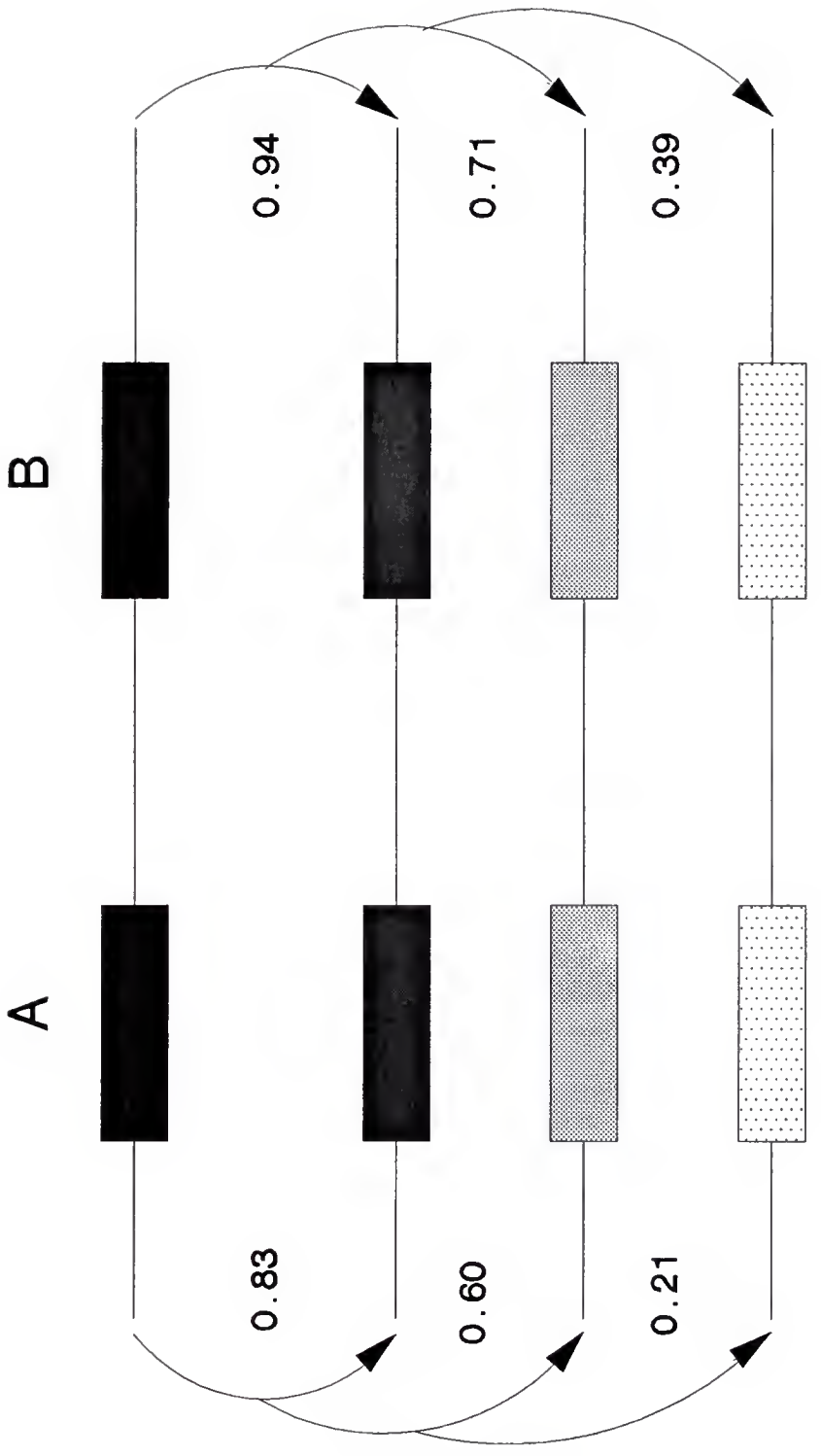
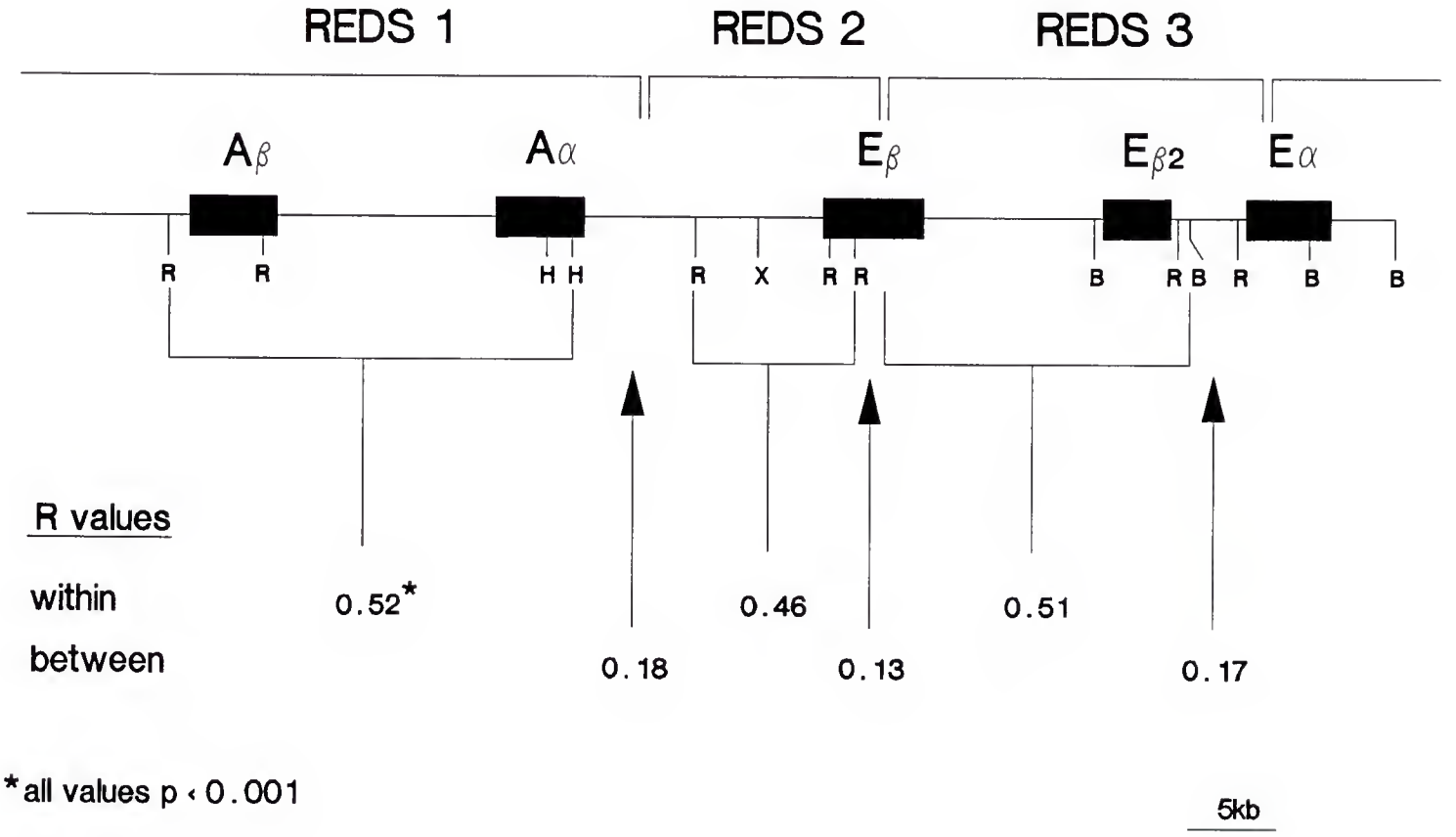

B
A
0.94
0.71
0.39
0.83
0.60
0.21

Table 4-17. Correlation analysis of loci across the I region.

locus	Aβ	Aα	I1	5'Eβ	3'Eβ	Eβ2	Eα
Aβ	-	0.52[a] .0001[b]	0.18 .0001	0.03 .4399	0.09 .0277	0.02 .7009	-0.11 .0069
Aα		-	0.18 .0001	-0.03 .4449	0.03 .4902	-0.04 .3418	-0.11 .0059
I1			-	0.46 .0001	-0.09 .0170	-0.18 .0001	-0.21 .0001
5'Eβ				-	0.13 .0006	-0.11 .0036	-0.19 .0001
3'Eβ					-	0.51 .0001	-0.04 .3557
Eβ2						-	0.25 .0001

[a]The correlation coefficient, R.
[b]The probability of R under the Ho; Rho=0.

Figure 4-7. Correlation of neighboring loci across the I region. Correlation coefficients (R values) are given at the bottom for loci within a REDS and for loci separated by recombination. As the strength of the correlation increases, the R values approach 1.0. All values were significant to the 0.0001 confidence level.


REDS 1
REDS 2
REDS 3
Aβ
Aα
Eβ
Eβ2
Eα
R
R
H H
R
X
R R
B
R B
R
B
B
R values
within
between
0.52*
0.18
0.46
0.13
0.51
0.17
*all values p < 0.001
5kb

line of values is for adjacent loci. These results are represented graphically in Figure 4-7. Loci within a REDS show a significant high correlation, on average 0.50, whereas loci flanking RHSs, therefore not within the same REDS, show very low correlation of around 0.15. These data are of twofold importance: (1) they confirm the location of the RHSs, and (2) they give insight into the relationships between the loci contained within a particular REDS. These relationships are twofold: (1) the genes within a REDS show a concerted evolution, i.e. they accumulate mutations in a coordinate manner, and (2) the genomic sequences that constitute a REDS tend to diverge as a single genetic unit.

Lineage Analysis of REDS

Because of the coordinate evolution of the genes within a REDS leading to the divergence of these genomic segments as a single genetic unit, they were analyzed as if they were a single gene. F values were calculated for each of the three REDS for all the haplotypes examined. These values were used to examine the lineage relationship between REDS much the same way as for the individual loci. Table 4-18 shows the results of this analysis. For REDS1, twenty-three of the twenty-eight haplotypes fall into three major evolutionary lineages which leaves only five unique forms of REDS1. Haplotypes within a particular REDS lineage show an average F value of 0.80 while haplotypes amoung separate

Table 4-18. I region diversity as defined by REDS.

REDS	lineages	F values		p value
		within	between	
1	11	0.80 ± 0.02	0.39 ± 0.01	<< 0.001
2	10	0.83 ± 0.04	0.36 ± 0.02	<< 0.001
3	4	0.80 ± 0.01	0.69 ± 0.01	< 0.001

lineages have F values less than 0.39. Recombinationally depressed segment 2 shows an analogous pattern where twenty-six of the strains are divided amoung four major evolutionary lineages with an F value of 0.83 within a particular lineage, and 0.36 between lineages. Recombinationally depressed segment 3, in the conserved tract of the I region, shows a lower number of more related forms analogous to that seen for the individual loci within this tract. There are only two forms of REDS3 which share 69% of their restriction fragments.

Haplotype Characterization of t Forms of Chromosome 17

The same analyses done in the wild type mice were also carried out on t-bearing mice. There are only seven t-bearing mouse strains in this analysis as opposed to the twenty-eight strains of wild type mice examined above (refer to Table 3-1). The results of the allele and F value analyses are shown in Figure 4-8. Although the allele and cluster number are lower for the t haplotypes, they correlate well with what is seen for the wild type mice. The F values are very similar, if not identical, to those calculated in the wild type mice. The I region of the t haplotypes is also divided into conserved and polymorphic tracts as seen with the wild type strains.

There are a large number of wild type alleles seen in the t haplotypes. A detailed analysis of the RFLP alleles

Figure 4-8. Characterization of the alleles for the seven probes for mice of the t haplotypes. Probe location, number of RFLP alleles for each probe, and average measure of diversity (F) as calculated are shown at the bottom.

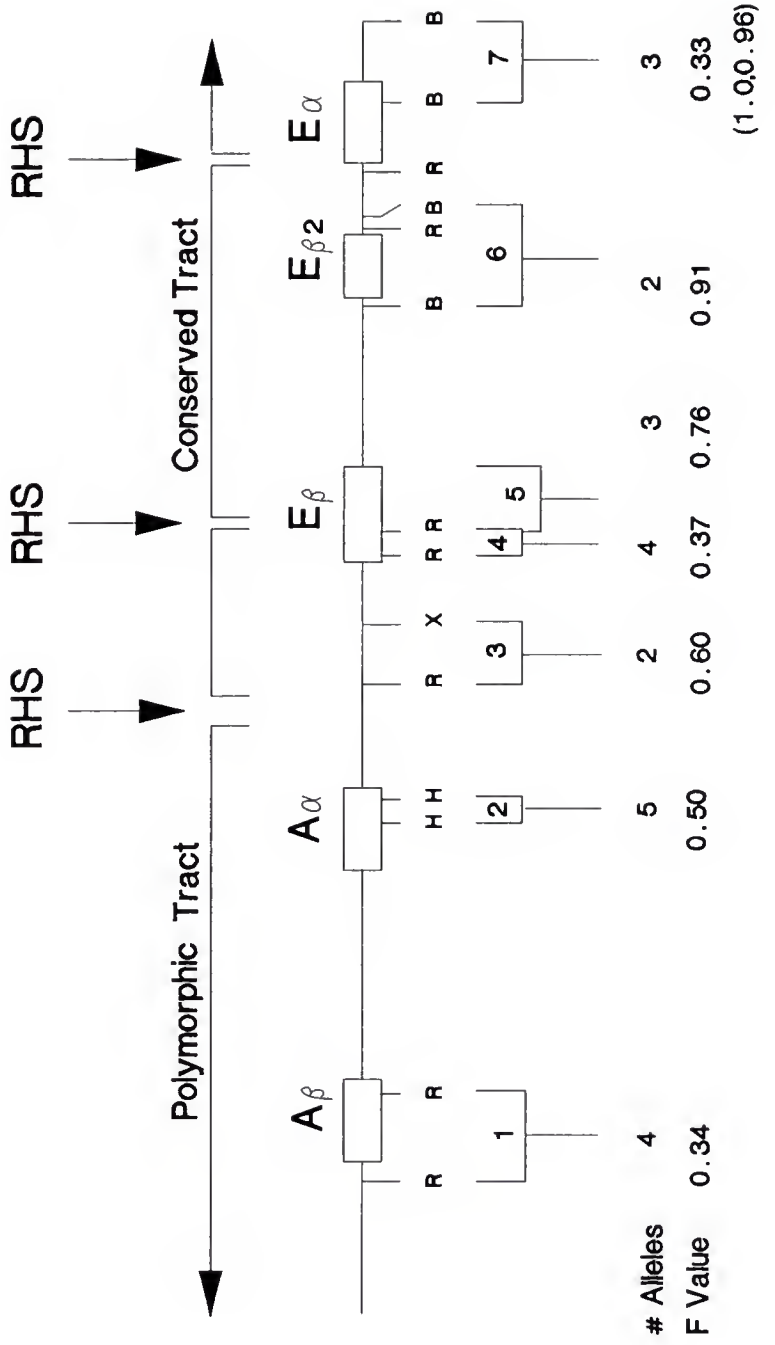

RHS
RHS
RHS
Conserved Tract
Polymorphic Tract
Aβ
Aα
Eβ
Eβ2
Eα
R R
H H
R X
R R
B RB R B
B
1
2
3
4
5
6
7
Alleles 4 5 2 4 3 2 3
F Value 0.34 0.50 0.60 0.37 0.76 0.91 0.33
(1.0,0.96)
5kb

Table 4-19. RFLP sizes and allelic grouping of t haplotypes for Aα.

strain	Bam HI[a]	Bgl II	Eco RI	Pvu II	Sac I	identical allele
t6	5.2	4.8	12.0	7.0	8.0 4.5	d, tw12, tw71
tw5	5.4	6.5	12.0	7.5 2.0	13.7 6.6	*
tw8	5.4	4.8	12.0	4.0	13.7 6.6	e, tw32
tw75	5.4	5.2	12.0	7.0	8.0 6.0	a

[a]Values expressed in kilobases. *Unique to t haplotypes.

Table 4-20. RFLP sizes and allelic grouping of t haplotypes for I1.

strain	Bam HI[a]	Bgl II	Eco RI	Pvu II	Sac I	identical allele
t6	10.0	6.0 1.0	8.5 2.0	2.0	8.0	*, tw12, tw71
tw5	10.0	5.0 2.0	9.0 2.0	2.0	9.0	a, tw8, tw32, tw75

[a]Values expressed in kilobases. *Unique to t haplotypes.

Table 4-21. RFLP sizes and allelic grouging of t haplotypes for 5'Eβ.

strain	Bam HI[a]	Bgl II	Eco RI	Pvu II	Sac I	identical allele
t6	3.8	3.0	1.0	3.4	2.5	f
tw5	8.6	3.0	2.0	3.4	5.2	a, tw8, tw32
tw12	3.8	3.0	1.0	3.6	2.5	*, tw71
tw75	7.5	3.0	2.0	3.0	4.2	*

[a]Values expressed in kilobases. *Unique to t haplotypes.

Table 4-22. RFLP sizes and allelic grouping of t haplotypes for 3'Eβ.

strain	Bam HI[a]	Bgl II	Eco RI	Pvu II	Sac I	identical allele
t6	10.0	2.8	12.2	4.5 3.0	2.6 2.4	e, tw8, tw32
tw5	10.0	3.1	12.2	4.5 3.0	2.6 2.4	l, tw75
tw12	10.0	4.6	12.0	4.5 3.0	2.6 2.4	*, tw71

[a]Values expressed in kilobases. *Unique to t haplotypes.

Table 4-23. RFLP sizes and alleleic grouping for t haplotypes for Eβ2.

strain	Bam HI[a]	Bgl II	Eco RI	Pvu II	Sac I	identical allele
t6	4.5	4.0 1.0	3.0 2.5	4.0 3.0	10.0 1.0	a, tw12, tw71
tw5	4.5	4.5 1.0 0.5	3.0 2.5	4.0 3.0	10.0 1.0	*, tw8, tw32, tw75

[a]Values expressed in kilobases. *Unique to the t haplotypes.

Table 4-24. RFLP sizes and allelic grouping of t haplotypes for Eα.

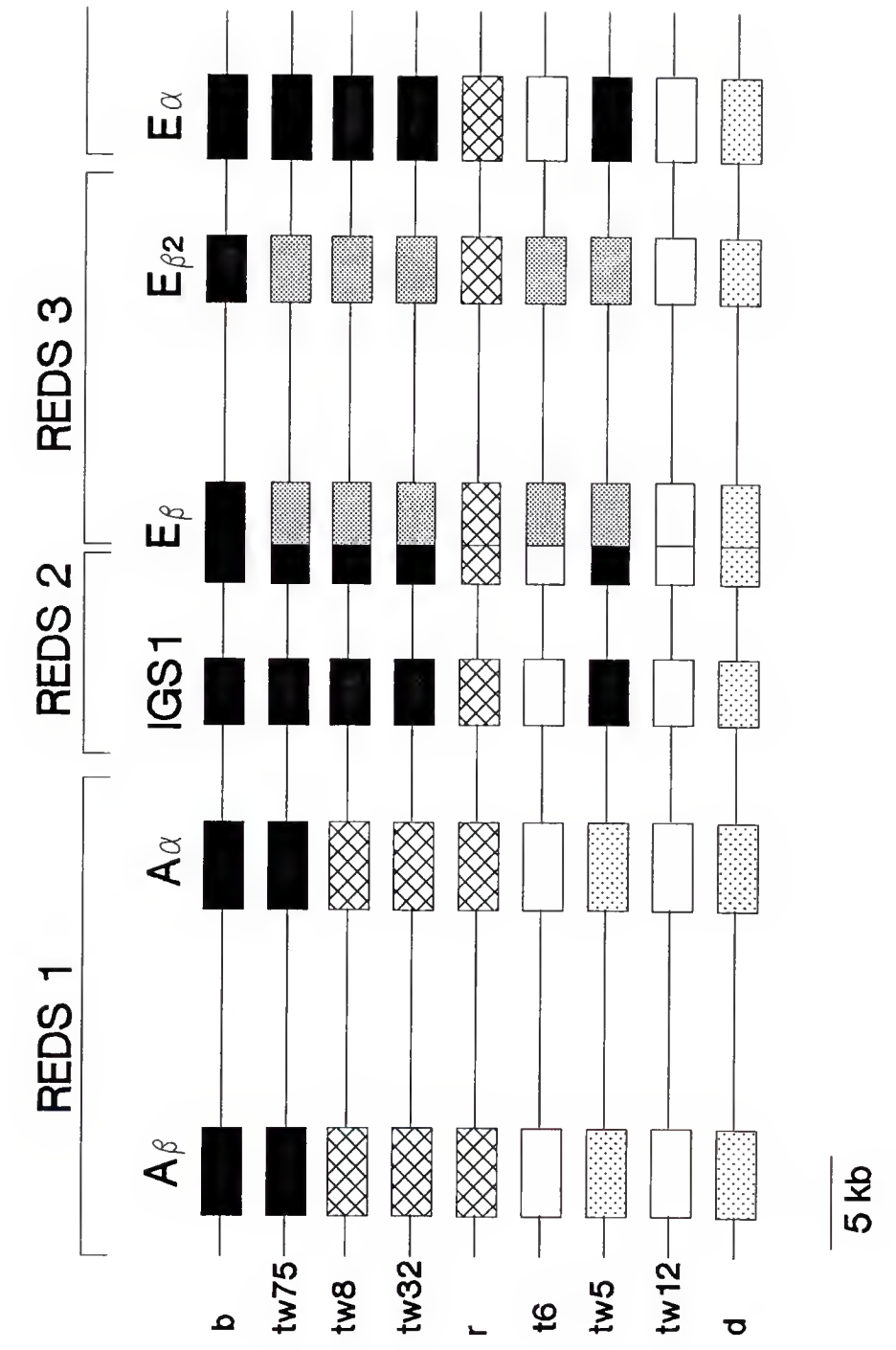
strain	Bam HI[a]	Bgl II	Eco RI	Pvu II	Sac I	identical allele
t6	6.5	6.0	8.5	7.5	6.0 1.2 0.7 0.4	*, bv1, tw12, tw71
tw5	9.0	5.5	8.0	7.0	8.0 2.5	*, dv1
tw8	6.0	5.5	8.0	7.0	8.0 2.5	d, tw32, tw75

[a]Values expressed in kilobases. *Unique to the t haplotypes.

for the t haplotypes is presented in Tables 4-19-24. In fact, there are only six alleles that are unique to the t haplotypes when all the loci are examined. Variants of wild type alleles that are unique to the t haplotype mice are seen but the majority of alleles are identical to the wild type haplotypes.

Another observation is that all the RHSs seen in the wild type haplotypes are active in the diversification of I region in t-bearing mice as well. If RHS activity is haplotype dependent, then only a subset of those seen in the wild type is expected to be seen in the t haplotypes. The presence of these RHSs and the influences they have on t haplotype associated I-regions can be seen in Figure 4-9. Recombination at the Aα-I1 site can be seen in tw8, tw32, and tw5. The Eβ RHS appears to be active in tw8, tw32, t6, and tw5. Recombination in the Eα RHS appears occur in tw8 and tw12. Although not as apparent at first, the division of the I-region into distinct REDS can be seen. By looking at the fill pattern associations within the expected REDS, one can see that stable associations exist between the loci. The t haplotypes show exactly the same organization and influences acting during haplotype generation as seen in wild type mice.

Figure 4-9. Recombination within the I region of t haplotype mice. Sites of recombination or RHSs are given for a representative collection of the t haplotype mice as compared to several wild type mouse haplotypes.


REDS 1
REDS 2
REDS 3
Aβ
Aα
IGS1
Eβ
Eβ2
Eα
b
tw75
tw8
tw32
r
t6
tw5
tw12
d
5 kb

CHAPTER V
DISCUSSION

Polymorphism of Genes Within the I Region

Mutations in DNA structure can be of three different forms; point mutations, insertions and deletions. RFLP analyses are capable of detecting all three types of mutations, although they are most sensitive in detecting insertions and deletions. Point mutations or single base changes are detected only when a mutations of this types alters the recognition sequence of one of the restriction endonucleases used in the analysis. Therefore, RFLP analysis underestimates the extent of point mutations relative to insertions or deletions, which will be detected regardless of the enzyme used as long as the sites flanked the mutation.

Within the I region, three of the seven probes identify regions that contain insertions or deletions. The Aβ gene contains a retroposon insertion in the second intron of H-2^b or evolutionary group 2 alleles (McConnell et al. 1988) and an additional insertion adjacent to this retroposon in the H-2^k or group 3 alleles. The Eβ gene contains three different insertional/deletional events. The H-2^b alleles have an extra 50-100 bps in the second intron adjacent to the β2 exon while the H-2^d allele have an additional 200 bps

in the first intron. Alleles related to the H-2^p haplotype have a 1.0 kb deletion in the second intron relative to the other allelic lineages. The third probe which detects a deletion mutation is for the Eα gene and shows the well documented 650 bp deletion in the 5' portion of the gene encompassing the promoter region. Therefore, of the seventy-six total alleles for all the loci probed, only six alleles can be distinguished solely on the basis of a single gross mutational event based on the five restriction enzymes in combination with the probes used in this study.

These results suggest that the most predominant mechanism for generation of diversity at a locus is a single base change, although this type of analysis would underestimate the extent of point mutations. This is not to say that this is the only mechanism active in the diversification of the class II genes and neighboring DNA. Insertions by retroposon elements, such as in Aβ, have been well documented and smaller insertions or deletions due perhaps to unequal crossing over, such as seen in Eβ, which is known to contain a hotspot for recombination, have also been seen.

Because RFLP analyses survey mostly non-coding sequences such as introns and intergenic stretches, this may not be an accurate measure of what is occurring in the coding sequences. There is strong evidence for intragenic recombination or gene conversion in the diversification of class II genes (McConnell et al. 1988; Mengle-Gaw et al.

1984) which may account for the clustering of mutations in the exon when their sequences are examined. Once more alleles are sequenced for the class II genes, the true magnitude of the role of intragenic recombination may be elucidated.

Unequal crossing over has been proposed as an additional mechanism for the generation of diversity in the I region. Recombination within the I region is of rather high fidelity due to the low number of small insertional or deletion events seen. Therefore, unequal crossing over plays a minimal role in generating diversity based on both the low incidence around RHSs, and the low incidence in exon sequences.

Grouping the alleles into evolutionary lineages based on gross mutational differences helps demonstrate that most polymorphisms within the I region are old events with only minor drift as seen by the low number of lineages at each locus within the polymorphic tract. This is also supported by examining the loci in the conserved tract, Eβ2 and Eα, where an extremely limited number of evolutionary forms are seen which show a low degree of diversity in comparison to loci within the variable tract.

It has been previously shown that, for Aβ, only a limited number of forms of this gene survived speciation (McConnell et al. 1988) and this, not surprisingly, appears to be the case for the entire I region. These evolutionary or progenitor haplotypes then underwent mutation, both point

mutation and homologous equal recombination, to generate the haplotypes prevalent in wild populations.

These same polymorphisms are observed for the t haplotypes in which there are a very large number of wild type alleles within these strains. If recombination is suppressed between the wild type and t forms of chromosome 17, the expected results would be to see RFLP alleles which are variants of the wild type but unique only to the t haplotype mice. The results obtained suggest that either the t forms of chromosome 17 are very recent mutations from the wild type, or, more probably, that there is much more recombination occurring between the wild type and the t haplotypes in the distal region than previously expected.

Recombination Within the I Region

Once markers for the different lineages of the loci within the I region were identified, the relationships between these loci could be examined. There is, in essence, one of two expected relationships between neighboring loci: linkage disequilibrium or recombination. If recombination was not occurring in the I region, the expected outcome would be strong associations between loci across the entire I region. If recombination is occurring and random, then when all haplotypes are examined together, there will be no linkage seen between any of the loci. These predictions represent the extremes and previous reports based on a

limited number of recombinant inbred strains have shown that neither appear to be the case (Steinmetz et al. 1982b, 1986, 1987; Lafuse et al. 1986). These studies were, however, limited in their scope. The data presented here is based on a broad, random sampling of wild derived haplotypes and is therfore more representative of wild mouse populations.

The results of this investigation reveal that recombination within the I region is not random and that it appears to be stringently regulated and restricted to specific sites. Of the sixteen recombinational events scored, six ocur in Eβ, four in Eα, and four between Aα and Eβ with two occurring elsewhere in the I region. Recombination rates are not high within the I region, being at 1 per 1000, given that the genetic map distance of the I region is 0.1 cM. Studies on the rate or frequency of recombination in the I region, based on screening for recombinant offspring of laboratory inbred mice, have shown that the numbers of recombinant offspring coincide with the predicted frequencies based on the genetic map distance (Steinmetz et al. 1987). The striking feature of recombination within the I region is its localization to recombinational hotspots, which are not randomly arranged within this genomic segment.

The other feature of recombination of importance within the I region is the haplotype specificity of the RHSs. The data presented in this dissertation show that the activity of the RHSs identified in the wild derived haplotypes in

this panel show a strong haplotype association. All recombination events in the Eβ RHS are in haplotypes sharing a 5' portion related to the H-2^b allele. All recombination events at the Aα to Eβ RHS have an I1 RFLP allele related to the H-2^b haplotype. Haplotype specificity for recombination in the Eα RHS is harder to determine due to the extreme relatedness of the alleles in this segment of the I region.

Whether or not this pattern of recombination is unique to the I region or is ubiquitous throughout the genome is difficult to assesses due to the low degree or the lack of polymorphism at most other loci. The only other system where recombination has been studied in a mammalian species is in the β-globin locus of the human. Recombination at this locus shows an analogous pattern of recombination to that in the H-2, i.e. being localized to a single site or RHS (Orkin and Kazazian 1984). When more polymorphic markers are identified for other loci in the genome, it will be of interest to see if this is a unique phenomenon or wide spread throughout the genome.

Within the t haplotypes, the same RHSs were identified. The predicted results would have been to see a subset of these RHSs based on the premise that the activity of the founder haplotype RHSs will dictate what is seen in the subsequent descendants of that strain. Because many of the same RFLP alleles are detected in the t haplotypes as seen in the wild type, there must be a higher rate of recombination between the two forms of chromosome 17 than

previously expected. Similarly, only a subset of the RHSs active in the diversification of the I region in the wild type would be expected to be seen in the t-bearing mice. The data presented here show the presence of all the active RHSs seen in the wild type and, therefore, further support the idea that recombination rates are higher between t and the wild type than expected.

The Influence of RHSs on the Evolution of I Region Haplotypes at the Genomic Level

As eluded to in the preceding section, lack of recombination can lead to linkage disequilibrium or vice versa. Therefore, what is seen in the relationship between two loci within a REDS? As shown in the last chapter, loci within a REDS show a strong linkage disequilibrium. These loci remain linked regardless of their haplotype. Recombination occurs at the RHSs leading to the shuffling of these linked loci between haplotypes.

These REDS represent discrete genetic units in which the genes have evolved in a coordinate manner. This suggests that during speciation, not just alleles of single genes survived, but REDS, represented by the selectable genes within them, were the units to survive speciation and any subsequent population bottlenecks.

With the pattern of evolution of I region haplotypes described, the question arises as to what forces dictate the

placement or activity of a RHS in different haplotypes. The site of a RHS is not random and the distance between any two RHSs can vary drastically. This is seen in the size differences between REDS where, for example, REDS1 encompasses 44 kb, REDS2 contains 19 kb, and REDS3 spans 30 kb. This shows that there is not a size constraint on the placement of RHSs within the I region, but that some other factor dictates whether recombination is allowed at a particular site.

Recombination, Selection, and Generation of I Region Haplotypes

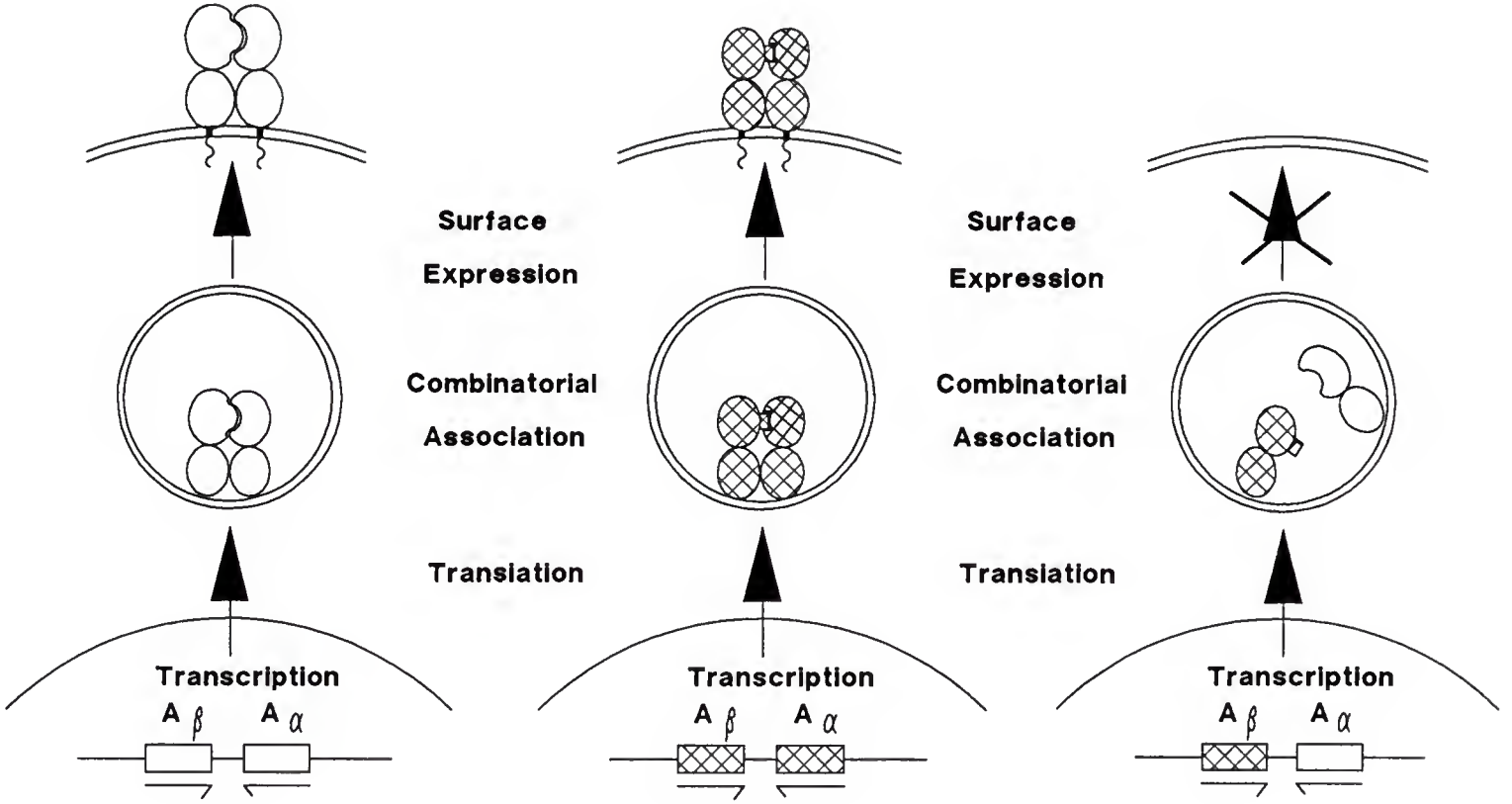
The real significance of this pattern of recombination within the I region can be appreciated when the functional role of the class II molecules are taken into account. Cell surface expression of these molecules depends on the association of the products of two genes: the α and β genes. When selection operates on two genes, the prediction is that there will be very low frequencies of recombination between these genes. This is exactly what is seen within REDS1 for the Aβ and Aα genes. The proper expression of the A molecule is dependent on the successful association of the two chains in the cytosol. By selecting against recombination between these two adjacent genes, the resulting linkage disequilibrium and subsequent co-adaptation of these two genes insures that the animal will

always be able to assemble and express an A molecule on its cell surface. This must be of extreme importance to the survival of the animal as reflected by the observation that there are no examples of mice lacking an A molecule.

Experimental evidence for the importance of the co-adaptation of these two genes has come out of the laboratory of Gemain et al. (1985) where it was shown that it is necessary to have haplotype matched $A\beta A\alpha$ pairs for the expression of an A molecule on the cell surface in transfection assays (Figure 5-1). Haplotype matched pairs represent an α and β chain from the same REDS1 lineage, whereas haplotype mismatched pairs represent the product of a recombinational event within this REDS, an event which is rarely ever seen.

The expression of the E molecule is under a different type of control which may reflect the lesser importance of the E molecule expression to the fitness and survival of the animal. It has been well documented that a relatively high frequency of wild mice do not express an E molecule and, therefore, the necessity of the E molecule for survival is questioned. Perhaps at some point earlier in the evolution of the mouse, the E molecule played an integral role in its survival, but this dependence on an E molecule apparently has been lost at some point prior to the divergence of M. m. domesticus as a separate subspecies. This is not to say that the $E\beta$ and $E\alpha$ genes are not co-adapted to some extent, but that the mechanism leading to the expression of the E

Figure 5-1. The importance of maintaining linkage disequilibrium for the expression of the A molecule. Chain association of the two class II molecule chains is shown. Haplotype mismatched $\alpha\beta$ pairs are transcribed and translated normally, yet fail to associate in the cytoplasm necessary for the expression of the molecule on the cell surface.


Surface
Expression
Combinatorial
Association
Translation
Transcription
A β
A α
Surface
Expression
Combinatorial
Association
Translation
Transcription
A β
A α
Transcription
A β
A α

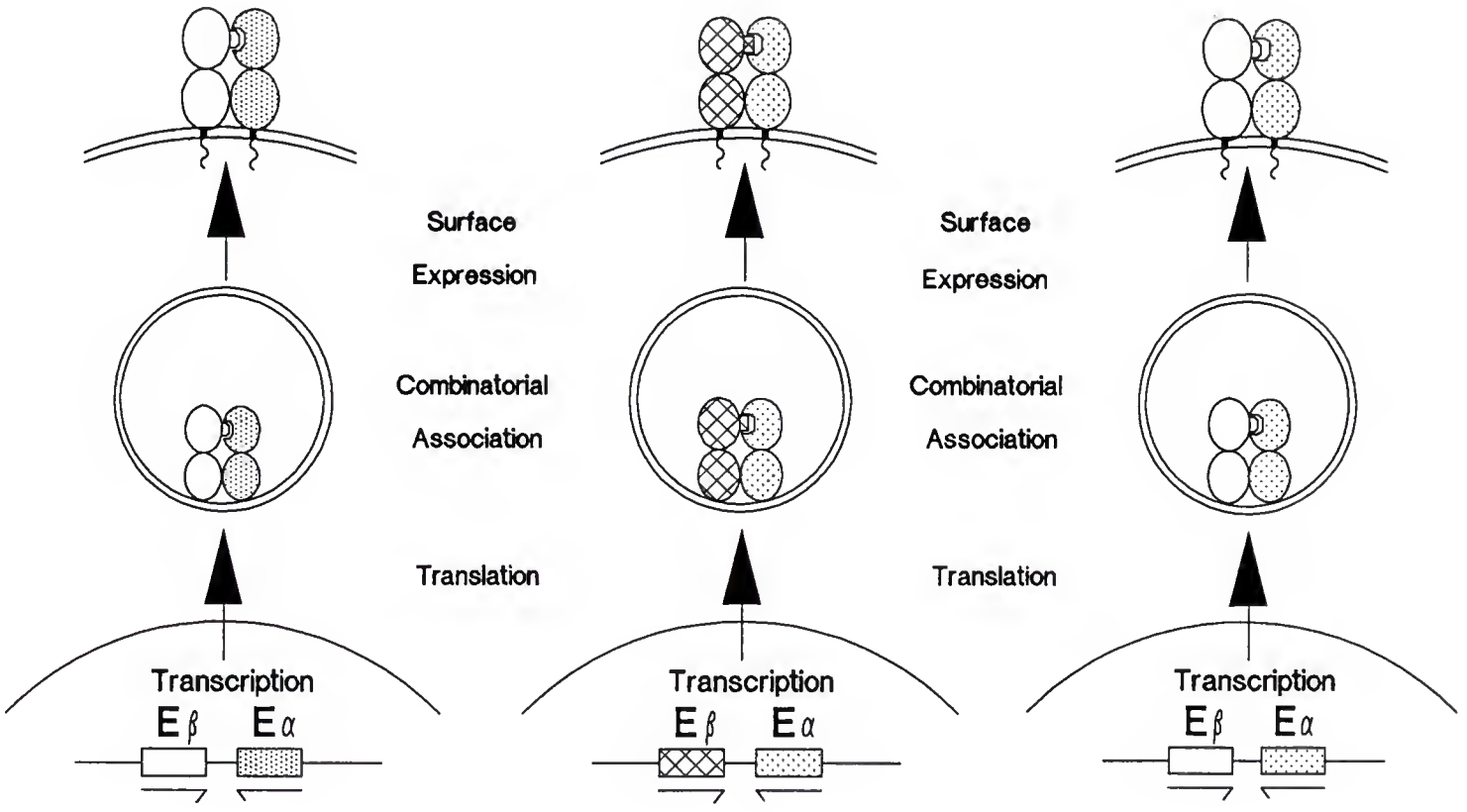
molecule differs significantly from that seen for the A molecule.

Whereas the α and β chains for the A molecule are encoded by genes within the same REDS, the E molecule α and β chains are encoded by genes on two separate REDS and are separated by two RHSs. Recombination can, therefore, shuffle different alleles for the α and β genes between haplotypes leading to a situation, if it was the A molecule, that will lead to the loss of expression. The majority of wild mice express an E molecule which suggests that there must be a mechanism which has evolved by which the E molecule can still be expressed regardless of recombination separating the two genes.

The Eα gene shows only two lineages, the expressed form and the deleted form. This monomorphic Eα chain has evolved so that it can associate with virtually any form of the Eβ chain. There also appears to be a very high frequency of a single allele of Eβ in this panel of mice, i.e. thirteen of the twenty-eight strains carry an H-2^b-like allele. This suggests that the maintenance of a single Eα morph which can associate well with one form of Eβ and to an adequate degree with all other forms of Eβ is the mechanism that has evolved to compensate for the inclusion of the α and β chain genes of this class II molecule in separate REDS (Figure 5-2).

There are, therefore, two mechanisms which have evolved to ensure the expression of a class II molecule on the surface of antigen presenting cells. The driving force

Figure 5-2. The importance of maintaining a monomorphic form of the Eα chain for the expression of the class II E molecule. Recombination is shown to have no effect on the association of the two chains due to the ability of the monomorphic α chain to associate with all β chains.


Surface
Expression
Combinatorial
Association
Translation
Transcription
Eβ Eα
Surface
Expression
Combinatorial
Association
Translation
Transcription
Eβ Eα
Transcription
Eβ Eα

behind the evolution of these two mechanisms appears to be selection either at the level of chain association or at the phenotypic level dictated by binding of foreign antigen for recognition by T lymphocytes. It is still not clear whether there is a positive selection for recombination at these RHSs, or a negative selection against recombination at sites which fall within REDS, or a combination of both. The most probable explanation is a combination of both where the placement of a RHS directs recombination to that site and thereby suppresses recombination for a distance flanking on either side. This paradox can be addressed, to some extent, by looking at the evolution of the RHSs within a panel of related species and subspecies of the genus Mus. This may shed some light on the question of whether recombination has always been confined to these RHSs, or, when viewed over an evolutionary course of 8 million years, recombination will appear to be random.

REFERENCES

Adorini, L., M.A. Harvey, A. Miller, and E.E. Sercarz. 1979. J. Exp. Med. 150:293.

Allen, P.M. 1987. Immunol. Today 8:270.

Allen, P.M., B.P. Babbitt, and E.R. Unanue. 1987. Immunol. Rev. 98:171.

Artzt, K., P. McCormick, and D. Bennett. 1982a Cell 28:463.

Artzt, K., H-S. Shin, and D. Bennett. 1982b Cell 28:471.

Artzt, K., H. Shin, D. Bennett, and A. Dimeo-Talento. 1985. J. Exp. Med. 162:93.

Babbitt, B.P., P.M. Allen, G. Matsueda, E. Haber, and E.R. Unanue. 1985. Nature 317:359.

Baltimore, D. 1981. Cell 24:592.

Baur, C.E., S.D. Hesse, J.F. Gardener, and R.I. Gumport. 1984. Cold Spring Harbor Symp. Quant. Biol. 49:699.

Bechtol, K.B., and M.F. Lyon. 1978. Immunogenetics 6:571.

Benacerraf, B., and M.E. Dorf. 1976. Cold Spring Harbor Symp. Quant. Biol. 41:405.

Benacerraf, B., and H.O. McDevitt. 1972. Science 175:273.

Bennett, D. 1975. Cell 6:144.

Bennett, D., K. Artzt, J. Cookingham, and C. Callo. 1979. Genet. Res. 33:269.

Benoist, C.O., D.J. Mathis, M.R. Kanter, V.E. Williams II, and H.O. McDevitt. 1983. Proc. Natl. Acad. Sci. USA 80:534.

Bjorkman, P.J., M.A. Saper, B. Samraoui, W.S. Bennett, J.L. Strominger, and D.C. Wiley. 1987. Nature 329:506.

Breakey, D.R. 1963. J. Mammal. 44:153.

Brown, J.H., T. Jardetzky, M.A. Saper, B. Samraoui, P.J. Bjorkman, and D.C. Wiley. 1988. Nature 332:845.

Buus, S., S. Colon, C. Smith, J. Freed, C. Miles, and H.M. Grey. 1986. Proc. Natl. Acad. Sci. USA 83:9368.

Buus, S., A. Sette, and H.M. Grey. 1987. Immunol. Rev. 98:115.

Calos, M.P., and J.H. Miller. 1980. Cell 20:579.

Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter. 1979. Biochem. 18:5294.

Choi, E., K. McIntyre, R.N. Germain, and S.G. Seidman. 1983. Science 221:283.

Condamine, H., J-L. Guenet, and F. Jacob. 1983. Genet. Res. 42:335.

Cullen, S.E., C.S. David, D.C. Shreffler, and S.G. Nathenson. 1974. Proc. Natl. Acad. Sci. USA 71:698.

Davis, M.M. 1985. Ann. Rev. Immunol. 3:537.

Dembic, Z., P.A. Singer, and J. Klein. 1984. EMBO J. 3:1647.

Dingler, L.C., R.T. Smith, L.J. Smith. 1987. Fed. Proc. Abstract #4315, 46:1057.

Dorf, M.E., and B. Benacerraf. 1975. Proc. Natl. Acad. Sci. USA 72:3671.

Duncan, W.R., and J. Klein. 1980. Immunogenetics 10:45.

Estess, P., A.B. Begovich, M. Koo, P.P. Jones, and H.O. McDevitt. 1986. Proc. Natl. Acad. Sci. USA 83:3594.

Falo, D., B. Benaceraff, and K.L. Rock. 1986. Proc. Natl. Acad. Sci. USA 83:6994.

Fathman, C.G., and M. Kimoto. 1981. Immunol. Rev. 54:57.

Ferris, S.D., R.D. Sage, E.M. Prager, U. Ritte, and A.C. Wilson. 1983. Genetics 105:681.

Ferris, S.D., R.D. Sage, and A.C. Wilson. 1982. Nature 295:163.

Figueroa, F., M. Golubic, D. Nizetic, and J. Klein. 1985. Proc. Natl. Acad. Sci. USA 82:2819.

Flaherty, L. 1980. The Role of the Major Histocompatibility Complex in Immunology. Dorf, M.E. ed., Garland, New York, p.33.

Flavell, R.A., and G. Widra. 1986. Science 233:437.

Gaisler, J. 1975. Rodents in Desert Enviroments. Prekash, I., and P.K. Ghosh, eds., Junk, The Hague, p.59.

Germain, R.N., D.M. Bentley, and H. Quill. 1985. Cell 43:233.

Gilfillan, S., M.E. Vadebomcoeur, D.M. Louie, and H.O. McDevitt. 1988. HLA/H-2 Wrkshp. 0:#125.

Gorer, P.A. 1936. Br. J. Exp. Pathol. 17:42.

Guillet, J.G., M.Z. Lai, T.J. Briner, S. Buus, A. Sette, H.M. Grey, J.A. Smith, and M.L. Gefter. 1987. Science 235:865.

Harland, P.S.E.G. 1958. Ann. Mag. Nat. Hist. 1:193.

Hassinger, J.D. 1973. Fieldiana, Zool. 60:1.

Herrmann, B., M. Bucan, P.E. Mains, A. Frischauf, L.M. Silver, and H. Lehrach. 1986. Cell 44:469.

Hood, L., M. Steinmetz, and R. Goodenow. 1982. Cell 28:685.

Hood, L., M. Steinmetz, and B. Malissen. 1983. Ann. Rev. Immunol. 1:529.

Hyldig-Nielsen, J.J., L. Schenning, U. Hammerling, E. Widmark, E. Helding, P. Lind, B. Servenius, T. Lund, R. Flavell, J.S. Lee, J. Trowsdale, P.H. Schreier, F. Zablitzky, D. Larhammar, P.A. Peterson, and L. Rask. 1983. Nucleic Acid Res. 11:5055.

Israel, A., A. Kimura, A. Fournier, M. Fellous, and P. Kourilsky. 1986. Nature 322:743.

Jeffreys, A.J., V. Wilson, and S.L. Thein. 1985. Nature 314:67.

Jones, J.K. Jr., and D.H. Johnson. 1965. Univ. Kans. Publ., Mus Nat. Hist. 16:357.

Jones, P.P., D.B. Murphy, and H.O. McDevitt. 1978. J. Exp. Med. 148:925.

Jones, P.P., D.B. Murphy, and H.O. McDevitt. 1981. Immunogenetics 12:321.

Kabat, E.A., T.T. Wu, and H. Bilofsky. 1979. U.S. Dept. of Health, Education, and Welfare, NIH Publication No. 80-2008, Bethesda.

Kaufman, J.A., C. Auffray, A.J. Korman, D.A. Shackelford, and J. Strominger. 1984. Cell 36:1.

King, D.P., and P.P. Jones. 1983. J. Immunol. 131:315.

Klein, J. 1973. International Symposium on Standardization of HL-A Reagents. Regamey, R.H., and J.V. Sparck, eds., Karger, Basel. p.251.

Klein, J. 1974. Ann. Rev. Genet. 8:63.

Klein, J. 1975. Biology of the Mouse Histocompatibility-2 Complex. Spinger-Verlag, New York.

Klein, J. 1980. Immunology 80 Fougereau, M., and J. Dausset. eds. Academic Press, London.

Klein, J., and F. Figueroa. 1981. Science 703:516.

Klein, J., F. Figueroa, and C.S. David. 1983a. Immunogenetics 17:553.

Klein, J., F. Figueroa, and Z. Nagy. 1983b. Ann. Rev. Immunol. 1:119.

Klein, J., M. Golubic, O. Budimir, R. Schopfer, M. Kasahara, and F. Figueroa. 1986. Curr. Top. Microbiol. Immunol. 127:239.

Klein, J., P. Sipos, and F. Figueroa. 1984. Genet. Res. 44:39.

Kobori, J.A., E. Strauss, K. Minard, and L. Hood. 1986. Science 234:173.

Kobori, J.A., A. Winoto, J. McNicholas, and L. Hood. 1984 J. Mol. Cell. Immunol. 1:125.

Lafuse, W.P., N. Berg, S. Savarirayan, and C.S. David. 1986. J. Exp. Med. 163:1518.

Lafuse, W.P., and C.S. David. 1986. Immunogenetics 24:352.

Larhammar, D., K. Gustafsson, L. Claesson, P. Bill, K. Wiman, L. Schenning, J. Sundelin, E. Widmark, P.A. Peterson, and L. Rask. 1982. Cell 30:153.

Larhammar, D., U. Hammerling, M. Denaro, T. Lund, R.A. Flavell, L. Rask, and P.A. Peterson. 1983a. Cell 34:179.

Larhammar, D., J.J. Hyldig-nielsen, B. Servenius, G. Andersson, L. Rask, and P.A. Peterson. 1983b. Proc. Natl. Acad. Sci. USA 80:7313.

Lieberman, R., W.E. Paul, W. Jr. Humphrey, and J.H. Stimphling. 1972. J. Exp. Med. 136:1231.

Little, C.C., and E.E. Tyzzer. 1916. J. Med. Res. 33:393.

Lyon, M.F. 1984. Cell 37:621.

Lyon, M.F., and R. Meredith. 1964a. Heredity 19:301.

Lyon, M.F., and R. Meredith. 1964b. Heredity 19:313.

Maddon, P.J., D.R. Littman, M. Godfrey, D.E. Maddon, L. Chess, and R. Axel. 1985. Cell 42:93.

Malissen, M., T. Hunkapiller, and L. Hood. 1983. Science 221:283.

Malone, R.E., D.K. Chattoraj, D.H. Faulds, M.M. Stahl, and F.W. Stahl. 1978. J. Mol. Biol. 121:473.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Cold Spring Harbor Laboratory, Cold Spring, New York.

Marshall, J.T. 1981. The Mouse in Biomedical Research, Vol. I. Foster, H.L., J.D. Small, and J.G. Fox, eds., Academic Press, Inc., New York, N.Y. p. 17.

Martin, W.J., P.H. Maurer, and B. Benacerraf. 1971. J. Immunol. 107:715.

Mathis, D.J., C.O. Benoist, V.E. Williams II, M.R. Kanter, and H.O. McDevitt. 1983. Cell 32:745.

McConnell, T.J., B. Darby, and E.K. Wakeland. 1986. J. Immunol. 136:3076.

McConnell, T.J., W.S. Talbot, R.A. McIndoe, and E.K. Wakeland. 1988. Nature 332:651.

McDevitt, H.O., and M. Sela. 1965. J. Exp. Med. 122:517.

McIntyre, K., and J. Siedman. 1984. Nature 308:551.

McNicholas, J., M. Steinmetz, T. Hunkapiller, P. Jones, and L. Hood. 1982. Science 218:1229.

Melchers, I., K. Rajewsky, and D.C. Shreffler. 1973. Eur. J. Immunol. 3:754.

Mellor, A.L., E.H. Weiss, K. Ramachandran, and R.A. Flavell. 1983. Nature 306:792.

Mengle-Gaw, L., S. Connor, H.O.McDevitt, and C.G. Fathman. 1984. J. Exp. Med. 160:1184.

Mengle-Gaw, L., and H.O. McDevitt. 1983. Proc. Natl. Acad. Sci. USA 80:7621.

Mengle-Gaw, L., and H.O. McDevitt. 1985. Ann. Rev. Immunol. 3:367.

Michaelson, J., E.A. Boyse, M. Chorney, L. Flaherty, I. Fleisner, U. Hammerling, C. Reinisch, R. Rosenson, and F-W. Shen. 1983. Transplantation Proc. 15:2033.

Moller, G. 1978. Immunol. Rev. 38:1.

Moller, G., ed. 1980. Immunol. Rev. 58.

Mond, J.J., S. Kessler, J. Kung, and F.D. Finkelman. 1981. J. Immunol. 127:881.

Monroe, J.G., and J.C. Cambier. 1983. J. Immunol. 130:626.

Murphy, D.B., L.A. Herzenberg, K. Okumura, and H.O. McDevitt. 1976. J. Exp. Med. 144:699.

Murphy, D.B., P.P. Jones, M.R. Loken, and H.O. McDevitt. 1980. Proc. Natl. Acad. Sci. USA 77:5404.

Nei, M., and W.H. Li. 1979. Proc. Natl. Acad. Sci. USA 76:5269.

Nizetic, D., F. Figueroa, and J. Klein. 1984. Immunogenetics 19:311.

Noelle, R., P.H. Krammer, J. Ohara, J.W. Uhr, and E.S. Vitetta. 1984. Proc. Natl. Acad. Sci. USA 81:6149.

Orkin, S.H., and H.H.Jr. Kazazian. 1984. Ann. Rev. Genet. 18:131.

Ozato, K., N. Mayer, and D.H. Sachs. 1980. J. Immunol. 124:533.

Parnes, J.R., and J.G. Seidman. 1982. Cell 29:661.

Paulnock-King, D., K.C. Sizer, Y.R. Freund, P.P. Jones, and J.R. Parnes. 1985. J. Immunol. 135:632.

Pearson, O.P. 1963. Ecology 44:540.

Phillip, U. 1938. J. Genet. 36:197.

Polla, B.S., A. Poljak, J. Ohara, W.E. Paul, and L.H. Glimcher. 1986. J. Immunol. 13:3337.

Radding, C.M. 1978. Ann. Rev. Biochem. 47:847.

Robertson, M. 1982. Nature 297:629.

Roehm, N.W., H.J. Leibson, A. Zlotnik, J. Kappler, P. Marrack, and J.C. Cambier. 1984. J. Exp. Med. 160:679.

Sage, R.D. 1981. The Mouse in Biomedical Research, Vol. I. Foster, H.L., J.D. Small, and J.G. Fox, eds., Academic Press, New York.

Saha, B.K., and S.E. Cullen. 1986a. J. Immunol. 136:1112.

Saha, B.K., and S.E. Cullen. 1986b. J. Immunol. 137:4004.

Saito, H., R.A. Maki, L.K. Clayton, S. Tonegawa. 1983. Proc. Natl. Acad. Sci. USA 80:5520.

Schwartz, R.H. 1985. Ann. Rev. Immunol. 3:237.

Sette, A., S. Buus, S. Colon, J.A. Smith, C. Mile, and H.M. Grey. 1987. Nature 328:395.

Sette, P.S., R.N. Moore, H.M. Johnson, and J.J. Oppenheim. 1982. J. Exp. Med. 152:1754.

Shimonokevitz, R., J. Kappler, P. Marrack, and H. Grey. 1983. J. Exp. Med. 158:303.

Shin, H-S., D. Bennett, and K. Artzt. 1984. Cell 39:573.

Shin, H-S., P. McCormick, K. Artzt, and D. Bennett. 1983. Cell 33:925.

Shiroishi, T., T. Sagai, and K. Moriwaki. 1982. Nature 300:370.

Silver, L.M. 1985. Ann. Rev. Genet. 19:197.

Singh, S.K., E.K. Wakeland, I. Vucak, Z.A. Nagy, and J. Klein. 1981. Immunogenetics 14:273.

Smith, G.R. 1983. Cell 34:709.

Smith, G.R., S.M. Kunes, D.W. Schulz, A. Taylor, and K.C. Triman. 1981. Cell 24:429.

Snyder, D.S., D.I. Beller, and E.R. Unanue. 1982. Nature 299:163.

Song, K-Y. 1985. Mol. Cell. Biol. 5:3331.

Soper, L.A., M.B. Stein, and H.C. Passmore. 1988. J. Immunol. 140:984.

Southern, E. 1980. Methods Enzymol. 69:152.

Steeg, P.S., R.N. Moore, H.M. Johnson, and J.J. Oppenheim. 1982. J. Exp. Med. 156:1780.

Steinman, R.M., N. Nogueira, M.D. Witmer, J.G. Tydings, and I.S. Mellman. 1980. J. Exp. Med. 152:1248.

Steinmetz, M., J.G. Frelinger, D. Fisher, T. Hunkapillar, D.Pereira, S.M. Weissman, H. Uehara, S. Nathenson, and L. Hood. 1982a. Cell 24:125.

Steinmetz, M., M. Malissen, L. Hood, A. Orn, R.A. Maki, G.R. Dastoornikoo, D. Stephan, E. Gibb, and R. Romaniuk. 1984. EMBO J. 3:2995.

Steinmetz, M., K. Minard, S. Horvath, J. McNicholas, J. Freelinger, C. Wake, E. Long, B. Mach, and L. Hood. 1982b. Nature 300:35.

Steinmetz, M., D. Stephan, and K. Fisher Lindahl. 1986. Cell 44:895.

Steinmetz, M., Y. Uematsu, and K. Fisher Lindahl. 1987. Trends in Genet. 3:7.

Stimphling, J.H., and T. Durham. 1972. J. Immunol. 108:947.

Streicher, H.C., I.J. Berrkower, M. Busch, F.R.N. Gurd, and J.A. Berzofsky. 1984. Proc. Natl. Acad. Sci. USA 81:6831.

Sukhatme, V.P., K.C. Sizer, A.C. Vollmer, T. Hunkapiller, and J.R. Parnes. 1985. Cell 40:591.

Tonegawa, S. 1983. Nature 302:575.

Uematsu, Y., K. Fisher Lindahl, and M. Steinmetz. 1986. EMBO J. 5:2123.

Unanue, E.R. 1983. Ann. Rev. Immunol. 2:395.

Wakeland, E.K., and B.R. Darby. 1983. J. Immunol. 131:3052.

Wakeland, E.K., B.R. Darby, and J.E. Coligan. 1985. J. Immunol. 131:2432.

Wakeland, E.K., and J. Klein. 1979a. Immunogenetics 8:27.

Wakeland, E.K., and J. Klein. 1979b. Immunogenetics 9:535.

Wakeland, E.K., and J. Klein. 1981. J. Immunol. 126:1734.

Wakeland, E.K., and J. Klein. 1983. J. Immunol. 130:1280.

Weiss, E., L. Golden, R. Zakut, A. Mellor, K. Fahrner, S. Kvist, and R.A. Flavell. 1983. EMBO J. 2:453.

Widera, G., and R.A. Flavell. 1985. Proc. Natl. Acad. Sci. USA 82:5500.

Winking, H., and J.L. Guenet. 1978. Mouse News Lett. 59:33.

Zaleska-Rutczynska, Z., and J. Klein. 1977. J. Immunol. 119:6.

Zinkernagel, R.M. 1979. Ann. Rev. Microbiol. 33:201.

BIOGRAPHICAL SKETCH

Roy William Tarnuzzer was born on June 29, 1960, in Orlando, Florida. There he grew up with his two brothers in suburban to rural surroundings. He graduated from Oak Ridge High School in 1978, and then attended the University of Central Florida in Orlando, Florida, where he earned his Bachelor of Science in microbiology in 1982. After toiling as a laboratory technician for six months, he started the graduate program in the Department of Pathology and Laboratory Medicine at the University of Florida in 1983. He received his Doctor of Philosophy degree from the Department of Pathology and Laboratory Medicine at the University of Florida in 1988.

I certify that I have read this study and that in my opinion it conforms to the acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Edward K. Wakeland
Edward K. Wakeland, Chairman
Associate Professor of
Pathology and Laboratory
Medicine

I certify that I have read this study and that in my opinion it conforms to the acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ammon B. Peck
Associate Professor of
Pathology and Laboratory
Medicine

I certify that I have read this study and that in my opinion it conforms to the acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Harry Ostrer
Assistant Professor of
Biochemistry and Molecular
Biology

I certify that I have read this study and that in my opinion it conforms to the acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William E. Winter
Assistant Professor of
Pathology and Laboratory
Medicine

I certify that I have read this study and that in my opinion it conforms to the acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Linda J. Smith
Assistant Professor of
Pathology and Laboratory
Medicine

This dissertation was submitted to the Graduate Faculty of the College of Medicine and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 1988

Dean, College of Medicine

Dean, Graduate School